1973 ANNUAL

PHILADELPHIA, PENNSYLVANIA JANUARY 23, 24, 25, 1973

PROCEEDINGS
1973 ANNUAL
RELIABILITY AND MAINTAINABILITY
SYMPOSIUM

PHILADELPHIA, PENNSYLVANIA JANUARY 23, 24, 25, 1973

TABLE OF CONTENTS

*Not available for publication.

*Not Available for Publication

*Not Available for Publication

*Not Available for Publication

AUTHOR INDEX

A CONSISTENT SHAPE PARAMETER ESTIMATOR FOR THE WEIBULL DISTRIBUTION

G. Arthur Mihram
University of Pennsylvania

In this paper is presented a new estimator for the shape parameter of the Weibull distribution. The estimator, being determined from the ratio of the sample arithmetic mean to the sample geometric mean, requires a complete random sample, yet is independent of the scale parameter. Moments of a function of the estimator are delineated explicitly; representative distributions associated with the estimator are determined by means of a Monte Carlo analysis. Comparisons, both with the maximum likelihood estimator and with the estimator based on the method of moments for logarithmic transformations of Weibull variates, conclude the paper.

1. Introductory Remarks

A positive-valued random variable, X, is said to have the (two-parameter) Weibull distribution [14] if its probability density function is given by

$$f(x;\ a,b) = \begin{cases} b.x^{b-1}e^{-(x/a)^b}/a^b, & x > 0 \\ 0 & , x \le 0, \end{cases} \qquad (1)$$

where the parameters a and b are each assumed to be positive throughout this exposition. The parameter, a, scales the random variables in the sense that, for $k > 0$,

$$Y = k\bullet X \sim f(y;\ ka,\ b),$$

where the symbol $\sim$ is read "is distributed according to", so that $Y_a = (X/a) \sim f(y\ ;\ 1,\ b)$, independently of a; whereas, the second parameter shapes the density, in the sense that the behaviour of the density (1) at the ordinate axis is dependent upon the relationship of the value of this parameter to unity, as depicted by Stacy and Mihram [12] in their Figure 1.

The family of densities (1) is also preserved under power transformations of the form $Z = X^t$; viz., for $t > 0$,

$$Z = X^t \sim f(z;\ a^t,\ b/t).$$

Consequently, the variate, $S = X^b \sim f(s;\ a^b, 1)$, which, upon reference to Equation (1), may be seen to be the exponential density of mean a^b. A standardized Weibull variate is then defined as an exponentially distributed random variable of unit mean; viz., $S_a = (X/a)^b \sim f(s;\ 1,\ 1)$.

Moments of Weibull variates are given by the general formula, provided that $t > -b$:

$$E[X^t] = a^t\ \Gamma(1 + t/b), \qquad (3)$$

where $\Gamma(v) = \int_0^\infty u^{v-1}e^{-u}du$ is the standard Gamma function of Euler of positive argument v. Thus, the mean of the Weibull variabe X is given by:

$$E(X) = a\ \Gamma(1 + 1/b); \qquad (4)$$

and its variance is given by

$$Var(X) = a^2\ \{\Gamma(1 + 2/b) - \Gamma^2(1 + 1/b)\}. \qquad (5)$$

More generally, for $t > -b/2$,

$$Var(X^t) = a^{2t}\ \{\Gamma(1 + 2t/b) - \Gamma^2(1 + t/b)\}. \qquad (6)$$

2. Shape Parameter Estimation

Given a complete random sample $x_1, x_2, \ldots, x_n$ from the density (1), one seeks to ascertain efficient estimators of the scale and shape parameters. A number of such joint parametric estimation schemes have been reviewed by the present author [10], an essential point therein being that generally one must first estimate the shape parameter b, then utilize this estimate to obtain a scale parameter estimate which is therefore functionally dependent upon the (nuisance) parameter, b. Only the maximum likelihood equations, when solved iteratively, and the graphical procedure of Kao [6], deviate from this pattern and provide estimates with absence of nuisance.

Consider the arithmetic mean of the sample

$$A = \sum_{i=1}^{n} \{x_i/n\}$$

and the geometric mean

$$G = \{\prod_{i=1}^{n} x_i\}^{1/n} = \prod_{i=1}^{n} \{x_i^{1/n}\},$$

the former having mean given by Equation (4), and variance by $n^{-1}Var\ (X)$ (cf: Equation (5)), whereas the latter has mean, derivable from repeated invocation of Equation (3):

$$E(G) = \{a^{1/n}\Gamma(1 + 1/nb)\}^n = a\Gamma^n(1 + 1/nb); \qquad (7)$$

and has variance

$$Var(G) = a^2\{\Gamma^n(1 + 2/nb) - \Gamma^{2n}(1 + 1/nb)\}.$$

Consequently, the ratio of the arithmetic mean to the geometric mean could be expected to have distribution independent of the scale parameter, a. As a matter of fact, the random variable

$$R \equiv A/G = \{\sum_{i=1}^{n} x_i\}/[n\{\prod_{j=1}^{n} x_j^{1/n}\}] \qquad (8)$$

is the equivalent of that in which $Y_{i(a)}$ is substituted herein for X_i, i = 1, 2, ..., n. The density function of R eludes precise description, but its mean may be determined by successive applications of the result in Equation (3); viz.,

$$E(R) = E\{\sum_{i=1}^{n} [X_i^{(n-1)/n} \prod_{\substack{j=1 \\ j\neq i}}^{n} X_j^{-1/n}]\}/n$$

$$= \{\sum_{i=1}^{n} E[X_i^{(n-1)/n} \prod_{\substack{j=1 \\ j\neq i}}^{n} X_j^{-1/n}]\}/n$$

$$= n\bullet a^{(n-1)/n}\Gamma[1 + (n-1)/(n\bullet b)]\Gamma^{n-1}[1 - 1/(nb)]\bullet a^{-(n-1)/n}/n,$$

or

$$E(R) \equiv \rho_n(b) = \Gamma[1 + (n-1)/(nb)]\, \Gamma^{n-1}[1 - 1/nb], \quad (9)$$

provided that $b > 1/n$.

The function $\rho_n(b)$ may be tabulated via computations of Equation (9), using standard tables of the Gamma function such as those provided by Abramowitz and Stegun [1]. The necessary computations for integral $n > 1$, were performed on the IBM 360 computer at the University of Pennsylvania, using linear interpolation for arguments of the Gamma function between those tabulated by Abramowitz and Stegun. Table 1 summarizes these computations for $n = 2, 5, 15, 25$, and ∞, with $b =$ 0.10 (0.10) 1.00 (0.20) 5.00.

The entries in Table 1 under the columnar heading "∞" may be computed by noting that

$$\lim \rho_n(b) = \lim \Gamma\left(1 + \frac{n-1}{nb}\right) \cdot \lim_{n\to\infty} \Gamma^{n-1}\left(1 - \frac{1}{nb}\right)$$

$$= \Gamma\left(1 + \lim_{n\to\infty}\left(\frac{n-1}{nb}\right)\right] \cdot \lim_{n\to\infty} \Gamma^n\left(1 - \frac{1}{nb}\right) / \left[\lim_{n\to\infty} \Gamma\left(1 - \frac{1}{nb}\right)\right]$$

$$= \frac{\Gamma(1+b^{-1}) \lim_{n\to\infty} \exp\{\ln \Gamma^n(1 - \frac{1}{nb})\}}{\Gamma(1)}$$

$$= \Gamma(1+b^{-1}) \cdot \exp\{\lim_{n\to\infty}[\ln \Gamma^n(1 - \frac{1}{nb}) - \ln \Gamma^n(1)]\},$$

since $\Gamma(v)$ is a continuous function of its argument with $\Gamma(1) = 1$. Therefore,

$$\lim_{n\to\infty} \rho_n(b) = \Gamma(1+b^{-1}) \exp\left\{\frac{-1}{b} \lim_{n\to\infty}[\{\ln \Gamma(1 - \frac{1}{nb}) - \ln \Gamma(1)\}/(-1/nb)]\right\}$$

or

$$\rho_\infty(b) = \Gamma(1 + b^{-1}) \cdot \exp\left\{\frac{-1}{b} \cdot \Psi(1)\right\}, \quad (10)$$

where $\Psi(v) \equiv d \ln \Gamma(v)/dv$, the digamma function, and where $\Psi(1) \cong -0.57722$, as given by Edwards [4, vol. II].

Of import is the observation that $\rho_\infty(b)$, as plotted in Figure 1, is the product of a pair of continuous and monotone decreasing functions of b, so that likewise is $\rho_\infty(b)$. Moreover,

$$\lim_{b\to 0^+} \rho_\infty(b) = +\infty,$$

and

$$\lim_{b\to\infty} \rho_\infty(b) = 1.$$

Furthermore, the statistic $R = A/G$ is positive and is bounded below by unity since the arithmetic mean A of a sample of positive variates is at least as great as the geometric mean G. (Cf: Kendall and Stuart, pp. 37-38, volume I, [7].) Suggested, therefore, as a Weibull shape parameter estimate is

$$\tilde{b} = \rho_\infty^{-1}(R), \quad (11)$$

where ρ_∞^{-1} is the inverse function of $\rho(b)$ and $R = A/G$ is the ratio of sample means, as given in Equation (8).

3. Statistical Properties of the Estimator

Thus, computation of the ratio R of the sample arithmetic and geometric means permits estimation of the Weibull shape parameter b:

(a) graphically, by means of Figure 1; or
(b) by means of linear or polynomial interpolation in tabulations of $\rho_\infty(b)$, such as those of Table 1.

From Equations (11), (10), and (8), the non-asymoptotic statistical properties of b remain obscure, though it would not appear likely that unbiassedness of the estimator should be such a property, the bias apparently depending both upon n and upon the underlying value of b.

The Cramer-Rao lower bound for estimates θ^* of some parametric function $\theta(b)$ is given by

$$\mathrm{Var}[\theta^*] \geq [\theta'(b)]^2 / E[-\partial^2 \ln L(X_1, \ldots, X_n; b)/\partial b^2],$$

where

$$L(x_1, x_2, \ldots, x_n; b) = \prod_{i=1}^{n} f(x_i; a, b)$$

is the likelihood function of the random sample. (See Wilks [15].) From Equation (1),

$$\frac{\partial^2 \ln L}{\partial b^2} = b^{-2} \{-n - \sum_{i=1}^{n} [X_i/a)^b] \cdot [\ln(X_i/a)^b]^2)\},$$

and hence

$$E[-\partial^2 \ln L/\partial b^2] = b^{-2}\{n + nE[S(\ln S)^2],$$

where $S = (X/a)^b \sim f(s; 1, 1) = e^{-s}$, $s > 0$.

Consequently,

$$E[-\partial^2 \ln L/\partial b^2] = (n/b^2)\, [\Gamma''(2)+1] \cong 1.8237\, n/b^2$$

where $\Gamma''(b) = \int_0^\infty u^{b-1} (\ln u)^2 e^{-u}\, du$, as given by Edwards [4]. It follows that, for θ^* an unbiassed estimate of $\theta(b)$,

$$\mathrm{Var}[\theta^*] \geq b^2[\theta'(b)]^2/n[\Gamma''(2) + 1]. \quad (12)$$

Now, the estimator $R = A/G$ of $\theta(b) \equiv \rho_\infty(b)$ has expectation as given by Equation (9), though the argument leading to Equation (10) provides the result that R is an asympototically unbiased estimate of $\rho_\infty(b)$. Furthermore,

$$E(R^2) = E[n^{-1} \sum_{i=1}^{n} X_i^{(n-1)/n} \prod_{\substack{j=1\\ j\neq i}}^{n} X_j^{-1/n}]^2$$

$$= n^{-2} E\{\sum_{i=1}^{n} X_i^{2(n-1)/n} \prod_{\substack{j=1\\ j\neq i}}^{n} X_j^{-2/n}\} +$$

$$+ n^{-2} E\{\sum_{\substack{i=1\\ i\neq k}}^{n} \sum_{k=1}^{n} X_i^{(n-1)/n} X_k^{(n-1)/n} X_i^{-1/n} \cdot$$

$$\cdot X_k^{-1/n} \prod_{\substack{m=1\\ m\neq i,\, m\neq k}}^{n} X_m^{-2/n}\}.$$

Repeated invocation of Equation (3) provides the result

$$E(R^2) = n^{-1}\{\Gamma^{n-2}[1-2/(nb)]\}.\{\Gamma[1+2(n-1)/(nb)]\cdot\Gamma[1-2/(nb)] + (n-1)\,\Gamma^2[1+(n-2)/(nb)]\},$$

provided that $b > (2/n)$.

The asymptotic variance of R may be found from this expression and Equation (9), noting that

$$\mathrm{Var}(R) = E(R^2) - \{E(R)\}^2.$$

One may express the second moment of R as

$$E(R^2) = n^{-1}C_n(b)\ \{\Gamma[1+2(n-1)/(nb)]\cdot\Gamma[1-2/(nb)] - \Gamma^2[1+(n-2)/(nb)]\} + C_n(b)\{\Gamma^2[1+(n-2)/(nb)]\}, \tag{13}$$

where

$$C_n(b) = \Gamma^{n-2}[1-2/(nb)].$$

In this form, one may note that, since the Gamma function is a continuous function of its (positive) argument,

$$\lim_{n\to\infty} C_n(b) = \frac{\lim_{n\to\infty}\{\exp[n\cdot\ln\ \Gamma[1-2/(nb)]\}}{\lim_{n\to\infty}\Gamma^2\ [1-2/(nb)]}$$

$$= 1\cdot\lim_{n\to\infty}\{\exp[(-2/b)\{\ln\ \Gamma[1-2/(nb)] - \ln\ \Gamma(1)]/(-2/nb)]\}$$

$$= \exp\ \{(-2/b)[\lim_{n\to\infty}\{\ln\Gamma[1-2/(nb)] - \ln\ \Gamma(1)\}/(-2/nb)]\}$$

$$= \exp -\{2\cdot\psi\ (1)/b\},$$

where $\psi(v)$ is as defined beneath Equation (10).

Considering in turn the braced quantities of Equation (13),

$$\lim_{n\to\infty}\{\Gamma[1+2(n-1)/(nb)]\cdot\Gamma[1-2/(nb)] - \Gamma^2\ [1+(n-2)/(nb)]\} = \Gamma\ [1+2b^{-1}] - \Gamma^2\ [1+b^{-1}],$$

and

$$\lim_{n\to\infty}\{\Gamma^2[1+(n-2)/(nb)]\} = \Gamma^2(1+b^{-1}),$$

then

$$\mathrm{Var}(R) \cong n^{-1}\ (\{\Gamma[1+2/b] - \Gamma^2[1+b^{-1}]\}\exp -\{2\cdot\Psi(1)/b\}) + [\Gamma^2[1+b^{-1}]\exp - \{2\Psi(1)/b\}] - \rho_\infty^2\ (b);$$

i.e., the asymptotic variance of $R = A/G$ is given by

$$\mathrm{Var}(R) \cong n^{-1}[a^{-2}\mathrm{Var}(X)]\exp-\{2\Psi(1)/b\}, \tag{14}$$

where Var(X) is as given in Equation (5). It follows that

$$\lim_{n\to\infty}\mathrm{Var}(R) = 0 = \lim_{n\to\infty}[E(R)-\rho_\infty(b)],$$

so that R is a consistent estimate of $\rho_\infty(b)$. (See Kendall and Stuart, pp. 3-4, volume II, [7].)

The estimator $\tilde{b}$, as given by Equation (11), is a consistent estimate of the Weibull shape parameter, since ρ_∞^{-1} is continuous and single-valued. (See Lukacs and Laha, [8].) The asymptotic variance of $\tilde{b}$ remains concealed, but a measure of the asymptotic efficiency of the estimator R of $\rho_\infty(b)$ is given by the ratio

$$\mathrm{leff}(R) \equiv \frac{\mathrm{Var}\ [n^{1/2}\ \rho^*]}{\mathrm{Var}\ [n^{1/2}R]},$$

where ρ^* is an hypothetical unbiassed estimate of $\rho_\infty(b)$ having variance given by the lower bound of Equation (12), with $\Theta(b) = \rho_\infty(b)$, (See Wilks, pp. 362-363 [15].), and where $\widetilde{\mathrm{Var}}[R]$ is the asymptotic variance given in Equation (14). The asymptotic efficiency of R is

$$\mathrm{leff}(R) = b^{-2}\cdot\{\Psi(1) - \Psi(1+b^{-1})\}^2/\{1.8237(CV)^2\}, \tag{15}$$

where

$$CV = [\{\Gamma(1+2b^{-1}) - \Gamma^2(1+b^{-1})\}\ /\ \Gamma^2(1+b^{-1})]^{1/2}$$

is the coefficient of variation of the density (1), as tabulated by Dubey [3].

Using the approximation

$$\mathrm{Var}\ [\tilde{b}] \cong \mathrm{Var}(R)\ \cdot\ \{d\rho_\infty^{-1}(R)/dR\ \Big|_{R=\rho_\infty(b)}\}^2$$

and $\theta(b) = b$ in Equation (12), an approximate espression for the asymptotic efficiency of $\tilde{b}$ becomes $\mathrm{leff}(\tilde{b}) \cong \mathrm{leff}(R)$. Table 2 provides these asymptotic efficiencies for b = 0.05(0.05) 1.00(0.10) 9.00.

4. Monte Carlo Analysis of the Estimation Procedure

In order to estimate the sampling distribution of the proposed estimates, $\tilde{b}$, 1000 random samples each of size n were generated from Weibull distributions having shape parameters b = 0.25(0.25) 2.5, the scale parameter value being immaterial. The sample sizes selected were n = 5(5)25, 50, 100. In each sampling experiment (i.e., with each specification of n and b), cumulative percentiles and other distributional properties were obtained for:

(a) R, the estimate of $\rho_\infty(b) \cong E[A/G]$;
(b) $\tilde{b}$, the corresponding estimate of the Weibull shape parameter; and
(c) $\tilde{c} = (\tilde{b}/b)$.

The statistical sampling experiments were conducted on the University of Pennsylvania's IBM 360 computer. For each stipulated size (n) and Weibull shape parameter (b), 1000n uniformly distributed random variables U_i were generated by means of the mixed congruential technique (See, e.g., Mihram [11]):

$$M_i = (1{,}000{,}000{,}005)M_{i-1}+(1{,}073{,}741{,}823)\quad (\mathrm{mod}\ ^{31}),$$

where

$$U_i = M_i/2^{31}\quad ,\quad i = 1,2,\ldots,\ 1000n.$$

In accordance with the Principia of Seeding (Mihram [11]), the seed values M_0 were selected randomly for each sampling experiment.

The necessary Weibull variates were then generated by means of the transformation

$$X_i = [-\ \ln U_i]^{1/b} \quad , \quad i = 1,2,\ldots,1000n$$

Concern regarding the possibility of excluding the extremely large values from the Weibull distribution was dispelled by noting that only values in excess of $\alpha \equiv [31 \ \ln 2]^{1/b}$ were precluded by the generator. The probability of such a large (or larger) value is given by

$$p = 1-F(\alpha;\ 1,b) = \int_\alpha^\infty b\chi^{b-1}\ e^{-\chi^b}\ d\chi = e^{-(\alpha)^b}$$
$$= \exp -\{\ \ln 2^{31}\} = 2^{-31} = 0.5 \times 10^{-9}.$$

Considering each sampling experiment as 1000n Bernoulli trials, each with probability of success p, then the Poisson approximation to the binomial distribution (See, e.g., Mihram [11] may be invoked to yield the approximate probability P that one or more extremely large values have been excluded from the experiment; viz.,

$$P = P[1 \text{ or more successes}] \cong n10^{-6}/2.$$

For sampling experiments herein reported, $n \leq 100$, so that

$$P < 10^{-4}/2.$$

In the context of the entire set of 70 sampling experiments conducted here*, the approximate probability, of one or more among the 2,250,000 Weibull variates' exceeding α, is only 1.1×10^{-3}, deemed by the author to be inconsequentially small.

The sampling properties of R were compared in each experiment with the results of Equations (9), (10), and (14). This served as a verification test both for the quality of the random number generator and for the accuracy of the computational routines. No discrepancies of statistical significance were detected in any of these verification tests.

The resulting sampling distribution of $\tilde{b}$ was obtained from each 1000 samples in the experiment corresponding to the specification of the "input parameters" n and b. Estimates of $\tilde{b}$ were obtained by linear interpolation in a precomputed table for $p_\infty(b)$, b = 0.03(0.01)10.00(0.05)35.00. However, in order to facilitate comparisons with the sampling distributions of the maximum likelihood estimator, as provided by Thoman, Bain, and Antle [13], and the moment estimators of Menon [9], only the sampling distributions of $\tilde{c} = \tilde{b}/b$ are presented. Table 3 provides the following properties of the sampling distribution of $\tilde{c}$, each distribution based on k = 1,000 samples of the given size (n):

1. <u>Ave$(\tilde{c})$</u>: The arithmetic mean of the observed sampling distribution. Its inverse would provide a bias correction factor which, if multiplied by $\tilde{b}$, would provide an approximately unbiassed estimate of b; however, for a given sample size, these bias correction factors depend upon the underlying and unknown value of b, so that Ave$(\tilde{c})$ is useful only in demonstrating the rate at which b becomes asymptotically unbiassed for b.

2. <u>S.D.$(\tilde{c})$</u>: The standard deviation of the observed sampling distribution:

$$[(999)^{-1}\sum_{j=1}^{1000}[\tilde{c}_j - \text{Ave}(\tilde{c})]^2]^{1/2},$$

where $\tilde{c}_j$ is the j^{th} estimate, j = 1,2,..., 1000.

3. <u>Min$(\tilde{c})$</u>: The smallest observed $\tilde{c}_j$, j = 1,2, ...,1000.

4. <u>Max$(\tilde{c})$</u>: The largest observed $\tilde{c}_j$, j = 1,2, ...,1000.

5. <u>Selected order statistics</u>: The values of $\tilde{c}_{(1000\gamma)}$, where $\tilde{c}_{(j)}$ is the j^{th} smallest among the $\tilde{c}_j$, j=1,2,...,1000, for γ=0.02, 0.05, 0.10, 0.15, 0.20, 0.25, 0.40, 0.50, 0.60, 0.75, 0.80, 0.85, 0.90, 0.95, 0.98.

5. <u>Comparisons with Alternative Estimators</u>

For a given sample size (n), Table 3 contains also previously published information regarding the corresponding statistical properties of the maximum likelihood estimator, $\hat{c} = \hat{b}/b$, and of Menon's moment estimator, $c^* = b^*/b$, based on ℓn X. Thoman, Bain, and Antle [13] note that both these estimators possess distributions which are independent of the underlying value of b and are asymptotically normal, though Johnson and Kotz [5, p. 256] report that earlier sampling studies had indicated that the approach to normality may be somewhat slow (0.8% bias remains for sample size 170). These results seem to be verified by the entries in Table 3.

Each of the histograms corresponding to the sampling distributions in Table 3 were compared with a normal distribution of mean Ave$(\tilde{c})$ and standard deviation S.D.$(\tilde{c})$ by means of a Chi-squared test of 16-1-2 = 13 degrees of freedom. In no instance, including all cases for which n = 100, was the hypothesis accepted (5% significance level) that the sampling distribution was the stated Gaussian form. The sampling distributions appear to be skewed to the right sufficiently to preclude their normality. Nonetheless, their asymptotic normality seemed apparent in that computed Chi-squared test statistics were observed to decrease monotonically as the sample size increased.

The search for efficient estimators of the Weibull shape parameter should be viewed in the context of the non-existence of a sufficient statistic for this parameter. It may be noted that the density (1) is not of the proper exponential type (See Kendall and Stuart, p. 26, volume 2, [7]) to possess a single sufficient statistic for this parameter. However, the density may be shown to satisfy the Wolfowitz regularity conditions, so that the Cramer-Rao lower bound of Equation (12) is applicable, though the absence of a sufficient statistic implies that it shall not likely be attained in estimating any non-trivial parametric function θ(b). (See Kendall and Stuart, pp. 8-9, 24-27, volume 2, [7]).

By comparison of the standard deviation of $\tilde{c}$ with that of Menon's estimator c*, it would appear that the present estimator is somewhat superior whenever n>15. This statement must be somewhat qualified in that the standard deviations of $\tilde{c}$ do not seem to be uniform in b, as is the case for the standard deviation of c*; however, unless b is

* Results in Table 3 exclude distributional properties of estimators for b=1.0, 1.5, 2.0, and 2.5, since these were insignificantly different from the neighbouring properties (b=0.75, 1.25, 1.75, 2.25).

expected to be extremely small or extremely large (i.e., outside the range of shape parameter estimates employed in this study), the present estimator would appear to be asymptotically about 20% more efficient than that of Menon [9]. (The assignment of an a priori distribution to the Weibull shape parameter might place the current discussion in its proper Bayesian context.)

Indeed, the comparison of the estimator $\tilde{c}$ with the distributional properties of the maximum likelihood estimator $\hat{c}$ would tend to support the notion that these be asymptotically equivalent estimators for "moderate" underlying shape parameters. For n > 15, comparison of the properties of $\hat{c}$ with those of $\tilde{c}$ whenever b=0.75 points to this result also.

In this regard, one may note the generalized ratio estimator

$$R_t = A_t/G_t \ , \qquad (16)$$

where

$$A_t = n^{-1} \sum_{i=1}^{n} X_i^{\,t}$$

and

$$G_t = \{ \prod_{i=1}^{n} X_i^t \}^{1/n} \ , \text{ for some } t > 0.$$

From Equation (2), it may be recalled that X_i^t has Weibull distribution of shape parameter $b_t = ^{-1}(b/t)$, so that the sampling distribution of

$$\tilde{b}_t = \rho_\infty^{-1} \ (R_t)$$

may be located in Table 3 by establishing the correspondence between the value of b there and that of (b/t) in the present context. Of possible interest is the observation that [See Table 2] the asymptotic variance of R_t is minimized when t=b/0.45; however, such a precise degree of a priori knowledge regarding the Weibull shape parameter would be inadmissible in the present parameter estimation context.

6. Summary

In this paper, a new shape parameter estimate for the Weibull distribution has been proposed. The estimation technique is independent of the Weibull scale parameter and involves the use of the inverse solution to a transcendental equation, the solution being probably best performed by interpolation within computed tabulations. The estimator is shown to be consistent and its asymptotic efficiency has been computed and tabulated.

Sampling experiments have been defined and their results presented; the estimator's distributional properties have been tabulated. The experiments reveal the applicability of the method as a useful alternative estimation technique, one not requiring the iterative solution of simultaneous, transcendental, equations such as those required by the method of maximum likelihood. Statistical properties of the estimates have been delineated and a comparison with the estimates of Menon [9] and of Thoman, Bain, and Antle [13] has been made.

One should note that the estimation of the Weibull scale parameter may proceed once the estimate, $\tilde{b}$, has been obtained. The reader is referred to the paper of Mihram [10] for an exposition of alternative scale parameter estimators.

In conclusion, the author wishes to express his appreciation for the efforts of Messrs. Steve Selcho and Edward Danielewicz, who conscientiously performed the computations and computer programming required in this study.

REFERENCES

1. Abramowitz, W. and I.A. Stegun (1964), Handbook of Mathematical Functions, Dover Publications, N.Y.

2. Bain, L.J. and E.E. Antle (1967), "Estimation of Parameters in the Weibull Distribution", Technometrics, 5, 175-182.

3. Dubey, S.D. (1967), "Some Simple Estimators for the Shape Parameter of the Weibull Laws", Naval Research Logistics Quarterly, 13, 489-512.

4. Edwards, Joseph (1954), A Treatise on the Integral Calculus (in two volumes), Chelsea Publishing Co., N.Y.

5. Johnson, N.L. and S. Kotz (1970), Distributions in Statistics: Continuous Univariate Distributions - 1, H. Mifflin, Boston.

6. Kao, J.H.K. (1959), "A Graphical Estimation of Mixed Weibull Parameters in Life-Testing of Electron Tubes", Technometrics, 389-407

7. Kendall, M.G. and A. Stuart (1958), The Advanced Theory of Statistics (in three volumes), Charles Griffin and Co., London.

8. Lukacs, E. and R.G. Laha (1964), Applications of Characteristic Functions, Charles Griffin, London.

9. Menon, M.V. (1963), "Estimation of the Shape and Scale Parameters for the Weibull Distribution", Technometrics, 5, 175-182.

10. Mihram, G. Arthur (1969), "Complete Sample Estimation Techniques for Reparameterized Weibull Distributions", I.E.E.E. Transactions on Reliability, R-18, 190-195.

11. Mihram, G. Arthur (1972), Simulation: Statistical Foundations and Methodology, Academic Press, N.Y.

12. Stacy, E.Webb and G. Arthur Mihram (1965), "Parameter Estimation for a Generalized Gamma Distribution", Technometrics, 7, 349-358.

13. Thoman, D.R., Lee J. Bain, and Charles E. Antle (1969), "Inferences on the Parameters of the Weibull Distribution", Technometrics, 11, 445-460.

14. Weibull, W. (1951), "A Statistical Distribution Function of Wide Applicability", Journal of Applied Mechanics, 8, 293-297.

15. Wilks, S.S. (1962), Mathematical Statistics, John Wiley, N.Y.

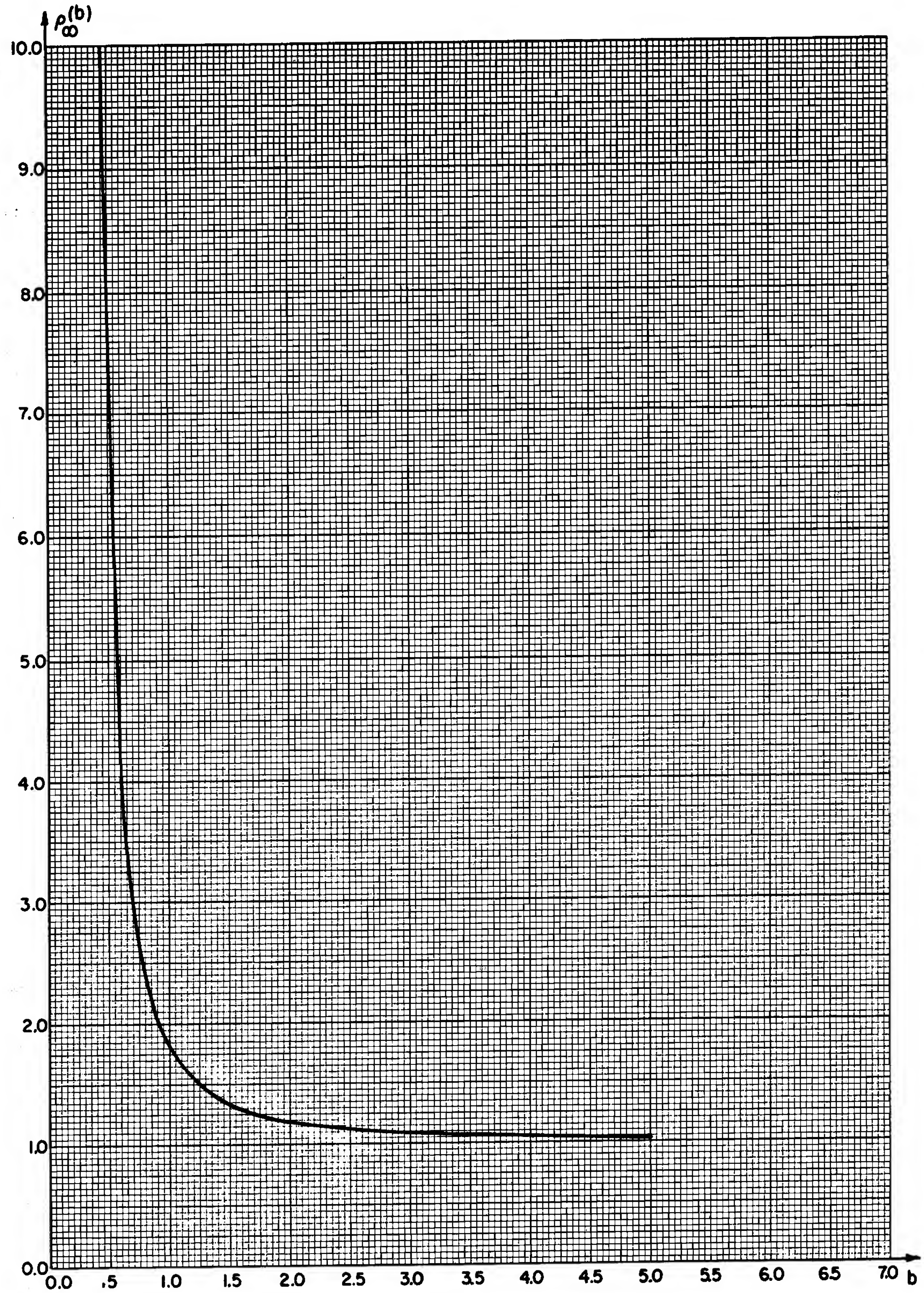

$\rho_\infty(b)$
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
0.0
.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
b

TABLE 1

TABULATION OF $\rho_n(b) = E[A/G]$

b	n = 2	5	15	25	∞
0.10	*	*	7.58×10^{11}	2.02×10^{10}	1.2×10^{9}
0.20	*	*	4,784.633	3,292.365	2,150.791
0.30	*	206.649	79.057	71.638	63.424
0.40	*	19.739	15.154	14.662	14.070
0.50	*	7.031	6.494	6.427	6.344
0.60	5.241	4.004	3.947	3.944	3.938
0.70	2.871	2.835	2.870	2.877	2.887
0.80	2.126	2.255	2.308	2.317	2.331
0.90	1.773	1.922	1.975	1.985	1.998
1.00	1.571	1.711	1.760	1.768	1.781
1.20	1.355	1.466	1.505	1.512	1.522
1.40	1.245	1.332	1.363	1.368	1.377
1.60	1.181	1.250	1.275	1.280	1.286
1.80	1.139	1.195	1.216	1.220	1.226
2.00	1.111	1.157	1.175	1.178	1.183
2.20	1.090	1.129	1.144	1.148	1.151
2.40	1.075	1.109	1.122	1.125	1.128
2.60	1.064	1.093	1.104	1.107	1.109
2.80	1.054	1.080	1.090	1.092	1.094
3.00	1.047	1.070	1.078	1.081	1.082
3.20	1.041	1.061	1.069	1.071	1.073
3.40	1.037	1.054	1.061	1.063	1.065
3.60	1.032	1.048	1.055	1.056	1.058
3.80	1.029	1.044	1.050−	1.050+	1.052
4.00	1.026	1.039	1.045	1.046	1.047
4.20	1.024	1.036	1.041	1.042	1.043
4.40	1.022	1.033	1.037	1.038	1.039
4.60	1.020	1.030	1.034	1.035	1.036
4.80	1.018	1.027	1.032−	1.032+	1.033
5.00	1.017	1.025	1.029	1.030	1.031

*$\rho_n(b)$ does not exist for $n \leq 1/b$.

TABLE 2

Asymptotic Efficiency of R = A/G

b	leff(R)	b	leff(R)	b	leff(R)	b	leff(R)	b	leff(R)
0.05	.0000	1.1	0.4829	3.1	0.0860	5.1	0.0334	7.1	0.0175
0.10	0.0025	1.2	0.4273	3.2	0.0811	5.2	0.0322	7.2	0.0171
0.15	0.0715	1.3	0.3798	3.3	0.0766	5.3	0.0310	7.3	0.0166
0.20	0.2847	1.4	0.3393	3.4	0.0724	5.4	0.0299	7.4	0.0162
0.25	0.5519	1.5	0.3045	3.5	0.0686	5.5	0.0289	7.5	0.0157
0.30	0.7710	1.6	0.2746	3.6	0.0650	5.6	0.0279	7.6	0.0153
0.35	0.9110	1.7	0.2486	3.7	0.0618	5.7	0.0269	7.7	0.0149
0.40	0.9809	1.8	0.2261	3.8	0.0587	5.8	0.0260	7.8	0.0146
0.45	1.0000	1.9	0.2063	3.9	0.0559	5.9	0.0252	7.9	0.0142
0.50	0.9870	2.0	0.1890	4.0	0.0533	6.0	0.0244	8.0	0.0139
0.55	0.9540	2.1	0.1736	4.1	0.0508	6.1	0.0236	8.1	0.0135
0.60	0.9101	2.2	0.1601	4.4	0.0485	6.2	0.0229	8.2	0.0132
0.65	0.8612	2.3	0.1480	4.3	0.0464	6.3	0.0222	8.3	0.0129
0.70	0.8107	2.4	0.1372	4.4	0.0444	6.4	0.0215	8.4	0.0126
0.75	0.7608	2.5	0.1275	4.5	0.0425	6.5	0.0208	8.5	0.0123
0.80	0.7128	2.6	0.1188	4.6	0.0408	6.6	0.0202	8.6	0.0120
0.85	0.6674	2.7	0.1109	4.7	0.0391	6.7	0.0196	8.7	0.0117
0.90	0.6248	2.8	0.1038	4.8	0.0376	6.8	0.0191	8.8	0.0115
0.95	0.5851	2.9	0.0973	4.9	0.0361	6.9	0.0185	8.9	0.0112
1.00	0.5483	3.0	0.0914	5.0	0.0347	7.0	0.0180	9.0	0.0110

TABLE 3

DISTRIBUTIONAL PROPERTIES OF $\tilde{c} = \tilde{b}/b$

BASED ON 1000 ESTIMATES

						Values of $\gamma \ni$ Prop $[\tilde{c} < \gamma]$ =														
n	b	Ave($\tilde{c}$)	S.D.($\tilde{c}$)	Min($\tilde{c}$)	Max($\tilde{c}$)	0.02	0.05	0.10	0.15	0.20	0.25	0.40	0.50	0.60	0.75	0.80	0.85	0.90	0.95	0.98
5	0.25	1.535	0.720	0.637	8.539	0.759	0.855	0.924	0.990	1.038	1.077	1.225	1.363	1.493	1.766	1.900	2.050	2.279	2.795	3.525
	0.50	1.432	0.708	0.511	7.840	0.658	0.733	0.818	0.889	0.937	0.994	1.129	1.235	1.381	1.673	1.808	1.960	2.213	2.708	3.365
	0.75	1.425	0.730	0.440	6.093	0.621	0.688	0.781	0.843	0.905	0.967	1.123	1.225	1.350	1.638	1.806	2.013	2.280	2.827	3.680
	1.25	1.433	0.798	0.438	7.580	0.582	0.670	0.754	0.818	0.873	0.931	1.095	1.242	1.398	1.667	1.815	1.994	2.247	2.879	3.602
	1.75	1.465	0.845	0.409	8.567	0.596	0.690	0.767	0.831	0.895	0.954	1.134	1.260	1.421	1.727	1.890	2.103	2.429	2.976	4.080
	2.25	1.474	0.855	0.458	7.096	0.556	0.649	0.748	0.811	0.870	0.931	1.113	1.249	1.396	1.713	1.892	2.126	2.449	3.058	4.005
	T/B/A	1.492	0.565			0.604	0.683	0.766			0.951	1.116	1.238	1.378	1.671	1.812	2.001	2.277	2.779	3.518
	Menon	1.370**	0.577																	
10	0.25	1.242	0.345	0.706	5.022	0.782	0.837	0.891	0.939	0.978	1.011	1.117	1.189	1.256	1.395	1.453	1.543	1.651	1.831	2.043
	0.50	1.177	0.333	0.584	3.047	0.694	0.776	0.832	0.868	0.896	0.936	1.035	1.110	1.198	1.339	1.407	1.465	1.613	1.837	2.053
	0.75	1.155	0.332	0.480	3.119	0.666	0.744	0.811	0.848	0.891	0.922	1.019	1.099	1.177	1.318	1.378	1.436	1.572	1.735	2.099
	1.25	1.170	0.354	0.572	3.040	0.658	0.712	0.782	0.835	0.877	0.920	1.026	1.053	1.199	1.360	1.430	1.523	1.628	1.796	2.163
	1.75	1.166	0.371	0.523	3.745	0.656	0.734	0.776	0.829	0.875	0.916	1.020	1.093	1.178	1.338	1.403	1.505	1.620	1.821	2.166
	2.25	1.164	0.375	0.499	3.763	0.642	0.706	0.776	0.835	0.877	0.912	1.018	1.079	1.161	1.343	1.405	1.491	1.624	1.885	2.124
	T/B/A	1.164	0.294			0.676	0.738	0.802			0.924	1.028	1.101	1.179	1.328	1.399	1.489	1.602	1.807	2.070
	Menon	1.129**	0.328																	

						Values of $\gamma \ni$ Prop $[\tilde{c} \le \gamma]$ =														
n	b	Ave($\tilde{c}$)	S.D.($\tilde{c}$)	Min($\tilde{c}$)	Max($\tilde{c}$)	0.02	0.05	0.10	0.15	0.20	0.25	0.40	0.50	0.60	0.75	0.80	0.85	0.90	0.95	0.98
15	0.25	1.173	0.249	0.691	2.468	0.790	0.837	0.892	0.934	0.969	0.997	1.077	1.137	1.196	1.309	1.345	1.404	1.483	1.613	1.813
	0.50	1.099	0.227	0.557	1.889	0.714	0.776	0.830	0.875	0.912	0.938	1.013	1.061	1.118	1.240	1.285	1.351	1.419	1.522	1.635
	0.75	1.107	0.259	0.614	2.735	0.713	0.765	0.821	0.863	0.897	0.939	1.019	1.071	1.129	1.232	1.274	1.319	1.411	1.584	1.788
	1.25	1.127	0.275	0.601	2.757	0.716	0.785	0.830	0.872	0.897	0.928	1.021	1.081	1.142	1.270	1.325	1.399	1.475	1.626	1.820
	1.75	1.105	0.266	0.556	2.273	0.677	0.739	0.803	0.846	0.884	0.918	1.003	1.066	1.130	1.244	1.301	1.362	1.461	1.623	1.738
	2.25	1.123	0.278	0.561	2.655	0.685	0.750	0.811	0.859	0.889	0.913	1.012	1.089	1.161	1.284	1.339	1.400	1.477	1.619	1.846
	T/B/A	1.102	0.226			0.716	0.770	0.823			0.925	1.008	1.064	1.124	1.234	1.284	1.349	1.427	1.564	1.732
	Menon		0.258*																	
20	0.25	1.137	0.200	0.720	2.223	0.820	0.856	0.904	0.940	0.965	0.994	1.063	1.113	1.165	1.249	1.284	1.342	1.404	1.480	1.595
	0.50	1.084	0.210	0.676	2.167	0.762	0.798	0.840	0.881	0.911	0.941	1.001	1.057	1.107	1.195	1.233	1.288	1.358	1.461	1.614
	0.75	1.078	0.214	0.623	2.424	0.748	0.808	0.846	0.874	0.902	0.922	0.991	1.044	1.096	1.188	1.236	1.289	1.350	1.445	1.615
	1.25	1.038	0.209	0.641	2.015	0.737	0.787	0.847	0.890	0.916	0.941	1.011	1.059	1.111	1.209	1.245	1.304	1.374	1.461	1.573
	1.75	1.074	0.222	0.598	2.556	0.726	0.772	0.825	0.859	0.887	0.913	0.994	1.047	1.097	1.196	1.243	1.294	1.364	1.444	1.590
	2.25	1.089	0.237	0.599	2.479	0.702	0.745	0.804	0.856	0.889	0.923	1.005	1.058	1.121	1.234	1.274	1.326	1.406	1.503	1.663
	T/B/A	1.072	0.190			0.743	0.791	0.838			0.929	1.000	1.047	1.097	1.188	1.228	1.281	1.343	1.449	1.579
	Menon	1.087**	0.224																	

TABLE 3

DISTRIBUTIONAL PROPERTIES OF $\tilde{c} = \tilde{b}/b$

BASED ON 1000 ESTIMATES

n	b	Ave($\tilde{c}$)	S.D.($\tilde{c}$)	Min($\tilde{c}$)	Max($\tilde{c}$)	0.02	0.05	0.10	0.15	0.20	0.25	0.40	0.50	0.60	0.75	0.80	0.85	0.90	0.95	0.98
						Values of $\gamma \ni$ Prop $[\tilde{c} \le \gamma]$ =														
25	0.25	1.117	0.188	0.677	2.360	0.800	0.852	0.905	0.940	0.962	0.984	1.051	1.090	1.138	1.214	1.257	1.301	1.366	1.444	1.565
	0.50	1.069	0.178	0.658	1.940	0.761	0.814	0.872	0.900	0.925	0.948	1.010	1.048	1.088	1.160	1.197	1.238	1.293	1.396	1.525
	0.75	1.056	0.180	0.590	1.928	0.750	0.796	0.840	0.877	0.908	0.932	0.994	1.038	1.088	1.167	1.196	1.226	1.283	1.372	1.485
	1.25	1.062	0.190	0.638	2.222	0.755	0.796	0.850	0.883	0.908	0.929	0.998	1.038	1.082	1.158	1.190	1.234	1.295	1.391	1.552
	1.75	1.054	0.191	0.642	1.900	0.730	0.794	0.831	0.860	0.887	0.915	0.981	1.028	1.080	1.175	1.211	1.262	1.309	1.395	1.488
	2.25	1.056	0.186	0.640	1.714	0.742	0.790	0.832	0.868	0.892	0.920	0.994	1.035	1.084	1.166	1.199	1.248	1.309	1.402	1.509
*	T/B/A	1.058	0.167			0.762	0.807	0.850			0.932	0.996	1.037	1.082	1.160	1.195	1.239	1.292	1.381	1.490
	Menon	1.064**	0.200																	
50	0.25	1.059	0.119	0.745	1.553	0.841	0.879	0.914	0.939	0.957	0.975	1.021	1.051	1.078	1.132	1.152	1.183	1.217	1.269	1.313
	0.50	1.037	0.118	0.771	1.592	0.834	0.868	0.899	0.924	0.938	0.952	0.996	1.021	1.052	1.106	1.123	1.152	1.193	1.263	1.303
	0.75	1.032	0.118	0.745	1.516	0.819	0.853	0.884	0.909	0.929	0.948	0.992	1.022	1.056	1.107	1.129	1.158	1.187	1.234	1.303
	1.25	1.029	0.120	0.702	1.468	0.817	0.852	0.884	0.910	0.928	0.946	0.989	1.015	1.046	1.102	1.128	1.156	1.191	1.241	1.309
	1.75	1.036	0.130	0.695	1.572	0.813	0.849	0.888	0.908	0.925	0.946	0.990	1.020	1.054	1.113	1.137	1.166	1.210	1.273	1.347
	2.25	1.036	0.136	0.715	1.577	0.806	0.845	0.879	0.900	0.920	0.938	0.990	1.020	1.053	1.110	1.134	1.173	1.215	1.284	1.370
	T/B/A	1.028	0.118			0.817	0.852	0.886			0.944	0.989	1.018	1.048	1.100	1.122	1.149	1.182	1.235	1.301
	Menon		0.145																	

n	b	Ave($\tilde{c}$)	S.D.($\tilde{c}$)	Min($\tilde{c}$)	Max($\tilde{c}$)	0.02	0.05	0.10	0.15	0.20	0.25	0.40	0.50	0.60	0.75	0.80	0.85	0.90	0.95	0.98
						Values of $\gamma \ni$ Prop $[\tilde{c} \le \gamma]$ =														
100	0.25	1.037	0.0845	0.749	1.329	0.879	0.909	0.933	0.950	0.963	0.978	1.011	1.032	1.052	1.093	1.109	1.126	1.148	1.182	1.230
	0.50	1.015	0.0816	0.795	1.352	0.860	0.884	0.914	0.932	0.948	0.961	0.993	1.013	1.031	1.065	1.082	1.093	1.117	1.155	1.199
	0.75	1.016	0.0833	0.796	1.302	0.859	0.895	0.919	0.933	0.943	0.956	0.989	1.009	1.030	1.069	1.081	1.104	1.125	1.156	1.201
	1.25	1.011	0.0867	0.779	1.302	0.860	0.884	0.902	0.915	0.934	0.948	0.983	1.006	1.026	1.071	1.087	1.109	1.126	1.156	1.210
	1.75	1.015	0.0886	0.784	1.394	0.850	0.879	0.906	0.925	0.940	0.952	0.990	1.008	1.031	1.068	1.087	1.103	1.129	1.166	1.214
	2.25	1.014	0.0935	0.757	1.409	0.848	0.870	0.901	0.920	0.937	0.951	0.986	1.007	1.033	1.071	1.087	1.111	1.134	1.172	1.230
	T/B/A	1.013	0.0774			0.861	0.888	0.916			0.956	0.988	1.009	1.029	1.065	1.079	1.096	1.116	1.150	1.192
	Menon		0.1050																	

* Obtained via interpolation from Thoman, Bain, Antle [13], Tables 1 and 6.

** From Bain and Antle [2].

SOME STATISTICAL TECHNIQUES USEFUL IN SYSTEM AGING STUDIES

Everett L. Welker, PhD.
TRW Systems
Redondo Beach, California

Summary

In order to maintain any system in a state of continuous alert with acceptable reliability over long periods of time, it is necessary to have a suitable understanding of the system aging characteristics. This is accomplished in aging surveillance programs which provide for monitoring pertinent system parameters over time to identify and measure characteristics of the degradation process. Regression methods have been widely applied in data analyses in aging studies. While these methods are often quite suitable, one frequently encounters degradation phenomena which are incompatible with the assumptions usually made in regression studies and for which traditional regression analyses can lead to serious errors. This paper discusses these problems and suggests some possible alternative non-parametric methods which avoid most of the restrictive assumptions. Aging trends, reliability functions and service life values are developed in probabilistic distribution form by using a Markov process approach. The paper discusses how these non-parametric methods can be combined with equipment condition monitoring to enhance the cost effectiveness of component change out policies based on specific equipment aging analyses.

Reliability trends are often checked by periodic functional tests over time. Even in the face of degradation in the survival probability, these tests may yield only successes due to the inability of the small samples to detect and measure the downward trend. The paper discusses a procedure for estimating values of the reliability function at times of test even when no failures have yet occurred. The method is non-parametric and it is compatible with the non-parametric reliability estimation procedure appropriate when the data includes failure events.

The paper includes actual illustrations of some of these techniques, one example relating to the problems associated with the use of regression analysis and another showing the estimation of a reliability function from test data which contains no failure events.

Regression Methods in Aging Studies

Reliability and aging surveillance programs have made extensive use of regression trend analysis including the associated confidence and tolerance limit computations. This is entirely natural since regression methods were developed to facilitate the description of phenomena like the time related degradation of hardware system properties. However, we are aware of the fact that traditional regression techniques involve assumptions and restrictions which impact on their capabilities to provide accurate and complete descriptions of system aging effects. Without involving any substantive loss in generality, we can identify the critical limitations of regression by a simplified description of the basic concepts and then by referring to the properties of linear regression analysis.

Consider a system characteristic, y, such as voltage, which must remain above some specification limit, say y_0. It is supposed that y degrades with age and we institute an aging surveillance program to monitor the system y values over time. We can conveniently think of such a program as applied to a group of individual systems such as the Titan or Minuteman missile force. The surveillance provides for measurement of y values for a sample of missiles on some periodic time schedule. Each such sample yields estimates of the distribution of y values in the force at the selected ages which we shall denote by x. In most cases, successive age samples are independent as opposed to repeated observations on the same items at different ages.

A schematic of the aging degradation phenomenon as it is usually generated in aging surveillance programs is given in Figure 1. In this figure we show a regression curve, the specification limit, and two densities at ages x_1 and x_2. Of course the collected data is used to derive the distributions and the regression curve, the curve being the age trend of the means of the distributions. The shaded area on the function at x_2 indicates the proportion of the force which has degraded below the acceptance specification limit. Service life is defined to be the age at which the proportion below specification becomes unacceptably high, with appropriate buffers for procurement lead time and other selected safety margins. Thus, the determination of service life involves management decision rules of behavior which are applied to the surveillance program data analysis. The resulting action is to replace or refurbish all items in the force when the service life age is reached. While this description is oversimplified, it still serves to define the basic concepts without significant distortion.

Consider now the most frequently used case, the one in which the regression function is assumed to be linear, say

$$y = A + Bx.$$

Data points (x_i, y_i) are obtained from the surveillance program and least squares or graphical fit by eye is used to obtain the estimated regression equation

$$y = a + bx.$$

It is usually assumed that the distributions about the regression have functions $f(y|x)$ which are Gaussian with constant variance independent of the age x. The next step is to compute a tolerance type prediction interval and use it in estimating useful service life in the following manner. Suppose management has selected a critical value, say ten percent, as the desired upper limit on the proportion of defectives in the force. For the one sided specification $\acute{y}>y_0$, we would compute the ten percent one sided prediction band below the regression related to the estimate of a future value of y associated with a given age x. Service life would terminate when the lower limit of this band reached the value $y = y_0$. If one chooses to remove the constant variance assumption, it is necessary to treat the data in homogeneous age groups and use Gaussian tolerance factors separately for

each age group. Figure 2 shows the results of a hypothetical regression computation, including the fitted regression line, the prediction limit for the constant variance assumption and tolerance limits for the non-constant variance case, the specification limit and the service life estimates, S_1 and S_2, for the two types of limits.

There are three basic problems with this type of analysis. First, the assumption of linearity may be invalid. A physics of failure analysis often suggests an asymptotic approach to a limiting value or perhaps a sharp drop in the parameter at a fairly well defined age. Second, the Gaussian assumption is not realistic for many parameters. The infinite range is impossible for such variables as voltage, tensile strength, thickness, resistance, etc. Third, the constant variance assumption is very often inconsistent with the facts. Many variables such as voltage and tensile strength will approach an asymptotic lower limit and will regularly show distributions about the regression line which approach a limiting variance near or equal to zero.

Serious errors can result from regression studies in which one or more of these problems are encountered. If degradation really isn't linear, extrapolating to describe ageout is almost sure to be wrong. If the Gaussian assumption is incorrect, then the wrong formula is used in computing the tolerance limits. When the variance changes over time, the regression line can describe the degradation rate only at the mean. This can be and usually is very serious. Specification limits are almost always far from the mean and degradation in the critical region can be very different from that indicated by the regression.

An Example to Illustrate Problems With Regression Methods

The failure of standard regression analysis to yield useful answers is illustrated by the following example. Although the data is unclassified, for proprietary reasons, it is inappropriate to identify the component. In an aging surveillance program, tensile strength of a laminate bond was tested on samples of systems at production and at later ages. It was believed that this parameter was the critical one with respect to potential failure. The following brief description of the analysis makes use of the actual test data.

The input data consists of two types of test results: (1) pounds of pull at failure and (2) pounds of pull at termination of test prior to failure. Each result of the second type is called a censored observation. It records a strength at failure at least as great as the observed pull at test termination as opposed to the exact strength at failure given by the first type of observation. The amount of test data is indicated in Table 1, Summary of Test Information.

Table 1. SUMMARY OF TEST INFORMATION

	Production		Retest at Age	
	Number	Percent	Number	Percent
Failures	136	15	236	42
Censorships	776	85	332	58
Total	912		568	

The extremely large amount of censorship reflected a termination of testing at an upper pull limit for many of the tests. Of course censorships also occurred at scattered lesser pull values. The first step in the analysis involved the computation of an observed reliability function by the method illustrated in Reliability Engineering, ARINC Research Corp., Prentice Hall, 1964. Since regression lines are the locus of distribution means, regression analysis requires the extrapolation of the observed survival functions beyond the upper pull limit at which censorship occurred to the mean value of the distribution. This is accomplished by grouping the data by age, computing observed survival functions for each group, and fitting Gaussian functions by probit type analysis, also described in the same ARINC book. Past studies on this tensile strength parameter have involved many different age groups. The current one was based on tests for two age groups, age zero and ages reasonably close to eight years. This provides a simple illustration of the points of interest. It oversimplifies the fitting of the regression line to the drawing of a line through two group means. The observed survival functions for the two ages and the fitted Gaussian function for the 8 year age group are shown in Figure 3.

The means and standard deviations for the fitted Gaussian functions were:

	Age 0	Age 8
Mean, $\hat{\mu}$	613	480
Standard Deviation, $\hat{\sigma}$	140	114.

The mean did decrease, indicating degradation. The standard deviation also decreased, thus precluding the assumption of constant variance over time. The impact of this changing variance is very significant.

Table 2. SUMMARY OF REGRESSION TYPE ANALYSIS

Group Age	$\hat{\mu} - 2\hat{\sigma}$	$\hat{\mu}$	$\hat{\mu} + 2\hat{\sigma}$
Zero	333	613	893
8 years	252	480	708
Difference	81	133	188
Annual Degradation Rate (lbs per year)	10	17	23

Table 2 shows the computation of average degradation rates at the mean and at the points $2\hat{\sigma}$ above and below the mean. The regression line reflects the rate of 17 pounds per year at the mean. The critical pull value, 250 pounds, is reached approximately at the $\hat{\mu} - 2\hat{\sigma}$ point at 8 years where the degradation rate is only 10 pounds per year. This illustrates the fact that service life forecasts cannot be made by using slopes of regression lines when the variance is changing. The regression slope is valid only at the mean and not at a specification limit at the low end of the distribution.

An appropriate non-parametric analysis is as follows. At age, for a pull of 250 pounds the observed survival probability is .9777. At age zero, this probability was associated with a pull of 329 pounds. Hence, in eight years, the degradation has been 329 minus 250 pounds, a total of 79 pounds or about 10 pounds per year. It is coincidence that this is the same as the $\hat{\mu} - 2\hat{\sigma}$ rate.

Further reference to this example will be made after discussing the following completely different approach.

A Non-parametric Method of Aging Analysis Using a Markov Process Approach

The weaknesses of regression analysis can be avoided by using a non-parametric method based on Markov process techniques. If the parameter y is really critical and if it is sensitive to age degradation, it should be selected as a precursor of ageout. That is, we are proposing that we use y measurements on specific systems to forecast specific system ageout, aging surveillance programs providing the data on degradation rates as a function of y. Markov methods are ideally suited for problems of this type.

The theory can be described in terms of a simple example. Let the variable y be represented by a discrete approximation consisting of four levels or states. Let p_{ij} be the probability that a system will transit from state i to state j in one unit of time. The matrix of p_{ij} values is the transition matrix of a Markov process,

$$T = \begin{Vmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{Vmatrix}$$

Let a_i denote the proportion of systems in state i at a specific point in time. Then the total population of systems at that time is described by the state vector

$$a = [a_1, a_2, a_3\ \ a_4].$$

After one unit of time, the state vector is transformed to a' by the relationship

$$a' = a\ T.$$

Repetition of this transformation gives the state vector after k units of time as

$$a\ T^k,$$

The state vectors for $k = 0,1,2, \ldots$ are in fact the fundamental description of the aging pattern of a collection of systems as illustrated in Figure 1. This theory is explained in standard texts.

A Hypothetical Example of the Non-parametric Method

The arithmetic has been carried out for a simple numerical four state example. The transition matrix is

$$T = \begin{Vmatrix} .5 & .5 & 0 & 0 \\ 0 & .5 & .5 & 0 \\ 0 & 0 & .1 & .9 \\ 0 & 0 & 0 & 1 \end{Vmatrix}$$

Let state one represent new equipment, state two and three being successively degraded states and let state four be failed equipment. The assumption of no self repair is indicated by the value $p_{44} = 1$. The unreliability function, U(t), the time to failure density, u(t) and the reliability function R(t) are developed and shown in Table 3 and plots of these functions are displayed in Figures 4 and 5. These functions are obtained by starting with the state vector $a = [1,0,0,0]$, all new equipments.

The regression function is the locus of the mean values of the state vectors over time. For this example, let the parameter values of the states be the state number. Column 8 of Table 3 lists the average state values which are the ordinates of the points on the regression curve. Figure 6 shows a plot of the regression curve and it also shows the histogram of one state vector, that for time equal to three. Note that the regression is nearly linear in the first portion of the curve, indicating the known risk of extrapolation error. It is important to observe that the regression function is completely defined by the transition matrix itself and no assumption need be made in advance about the form of the curve.

The Estimation of Service Life Using the Non-parametric Method

In this approach, service life is estimated simply from the conditional survival probabilities. Suppose a system has parameter value y_i. The conditional survival function, $R(t|y_i)$, is derived by a procedure like that illustrated in Table 1, using an initial state vector which assigns unity as the probability that $y = y_i$.

The curve derived in Table 1 is $R(t|y_1)$, the initial vector being (1,0,0,0). For $R(t|y_2)$, the initial vector is (0,1,0,0). Figure 7 shows plots of these functions.

For a system which has been in state y_i for time t_0, the probability of surviving for another t time units is

$$\frac{R(t + t_0|y_i)}{R(t_0|y_i)}$$

The service life is computed by solving for the time t at which this probability reaches an assigned critical level. For example, if a maximum of ten percent defective is to be accepted, time t would be determined by solving the equation

$$\frac{R(t + t_0|y_i)}{R(t_0|y_i)} = .9$$

At the time of transition into state y_i, the service life is estimated as the value of t satisfying the equation

$$R(t|y_i) = .9.$$

A Comparison of the Two Methods For Analyzing Aging Trends and Estimating Service Life

The regression method requires assumptions about the shape of the regression trend line or curve and the nature of the distributions about the regression trend. The most common assumptions are linearity and Gaussian distributions with time independent variances. None of these assumptions are needed for the non-parametric method described herein. It generates the regression curve and the distributions about the curve. The only critical assumption is that the parameter being measured is indeed a good precursor of failure and

ageout and that a first order Markov process is adequate.

The rule of behavior in the regression method for estimating service life is to discard all equipments at a specific age. At a ten percent defective level, this results in discarding equipments of which at least ninety percent are still good. The regression method works on a population rather than an individual basis, so it does not attempt to identify the "lemons". By contrast, the Markov process non-parametric approach classifies individuals on the basis of the critical age sensitive parameter and discards only those equipments which have failed or which have an unacceptably high risk of failure. That is it identifies and discards the bad ones but it does not throw out the good ones. The attractiveness of this cost saving procedure is obvious.

Predicting Survival Probabilities and Service Life From Tests With No Failures

This subject can be described most easily by an actual example, the part being unidentified for proprietary reasons. The U.S. Air Force experienced an unacceptably high failure rate from a particular expensive part. A vendor developed an entirely new one which was even more costly. Eight were provided for field test and they performed without failure as follows:

Number	Time of Failure Free Operation
4	196
2	235
2	237

The part appeared to be far superior to the older version and USAF wished to sign a procurement with an appropriate mean time to failure guarantee. Thus, it was necessary to estimate an MTTF from data with no failure events, the ultimate in censorship. The procedure which was used is as follows.

Suppose N items are placed on life test. If there is no censorship, the expected value of the reliability function at the time of the i th ordered failure is

$$R(t_i) = \frac{N - i + 1}{N + 1}$$

The extension to the case with censorship as developed by G. R. Herd is described in the previously referenced ARINC reliability book. Since we wish to view the N items on test as a sample from an infinite population, it is appropriate to use the concept of Yates' correction for continuity. Thus, we view the i th failure as really covering the range i - 0.5 to i + 0.5. For zero failures the range is from 0 to 0.5, thereby saying that zero failures really means not more than 0.5 failures. Combining these two concepts, we can say that for zero failures in N tests, we can place a lower bound on the estimate of reliability by taking i = 0.5. This gives

$$\hat{R}\,(t_i{=}0) > \hat{R}\,(t_i{=}0.5) = \frac{N + 0.5}{N + 1}\ ,$$

the fraction being a slightly conservative estimate of the value of the reliability in the zero failure case.

We recognize that unity is the maximum likelihood estimate of reliability when tests result in no failures. We are not comfortable with this value if we really believe that reliability is decreasing with age but that the small sample sizes have not been able to detect the suspected degradation. Our objective is merely to generate a reasonable estimate less than unity which is compatible with the data and the degradation hypothesis. The above theory appears to accomplish these objectives. Details of the arithmetic for the numerical example are given in Table 4.

The first two columns repeat the basic data. The third cumulates the n values, reflecting the assumption that operability at time t verifies operability for all previous time. The first reliability estimate is given directly from the formula as described above. The second estimate interprets the values in the fourth column as being conditional probabilities. Further research is needed on this point. It is believed that the second reliability estimates in the fifth column are unduly conservative. However, in the application made by USAF the conservative form was recommended and used. The last column lists the normal deviate corresponding to the reliabilities in column five. The Gaussian reliability function parameters are obtained by fitting a line to the three points, (z_i, t_i) as discussed above. For this example, the fitted line has the equation

$$t = 265 + 40\ z,$$

giving 265 as an estimate of the mean time to failure and 40 as an estimate of the standard deviation.

This method for handling zero failures is essentially consistent with the ARINC technique involving failure events. It is interesting that since the numerical analysis, some of the test items have failed and the estimate of the mean life still seems to be holding. Most of the alternative approaches address the zero failure case by an attempt to obtain a lower confidence bound of some type. Such a bound is not an adequate replacement for a point estimate. Thus far, analyses have justified an optimism that the proposed point estimate will prove to be quite suitable.

Conclusions

Traditional regression methods must rely on three basic assumptions:

1. The regression shape, usually taken as linear.
2. The distribution about the regression, usually taken to be Gaussian.
3. The variances of the distributions about the regression, often assumed to be constant over time.

The non-parametric method described herein eliminates all of these since it generates the regression curve and the distributions about the regression. It merely assumes that a first order Markov process is either correct or is a suitable approximation.

The service life estimates generate quite different rules of behavior. Regression analysis provides replacement based on an age criterion while the non-parametric method uses the age sensitive parameter on an individual basis. The age criterion is less cost-effective if parameter measurement procedures are not costly and are not destructive.

The regression method must be used very carefully when variance is time dependent. The degradation rate of the regression line is descriptive for the mean

values and this rate is not necessarily descriptive of degradation at values distant from the mean. Incorrect use of the regression rate can result in very serious and costly errors.

A method is described for estimating ovserved reliability functions from tests with no failure events. The method derives point estimates of survival probabilities in a manner which is consistent with the one in which failures do occur. The zero failure results appear to be reasonable, but further research in this area is indicated.

Table 3. COMPUTATIONS FOR THE NUMBERICAL EXAMPLE

Time (t)	State				Failure Density Function u(t)	Reliability Function R(t)	Average State Values
	a_1	a_2	a_3	a_4 Failure U(t)			
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0	1	0	0	0		1	1.00
1	.5	.5	0	0	0	1	1.50
2	.25	.5	.25	0	0	1	2.00
3	.125	.375	.275	.225	.225	.775	2.60
4	.0625	.250	.215	.4725	.2475	.5275	3.10
5	.03125	.15626	.1465	.6660	.1935	.3340	3.45
6	.015625	.09375	.092775	.79785	.13185	.20215	3.67
7	.0078125	.0546875	.0561525	.8813475	.0384675	.1186425	3.81
8	.0039062	.03125	.032959	.9318848	.0505372	.0681152	3.89
9	.0019531	.0175781	.0189209	.9615478	.0296631	.0384522	3.94
10	.0009766	.0097656	.0106812	.978577	.0170288	.0214233	3.97
11	.0004883	.0053711	.0059509	.9881897	.0096130	.0018103	3.98
t >11	0	0	0	1.000	0118103	0	4.00

Table 4. COMPUTATIONS FOR ZERO FAILURE EXAMPLE

Time	Number of Item	Cumulative n	Reliability Estimates - Two Forms		Normal Deviate
t	n	N	$\frac{N+.5}{N+1}$	R(t)	Z
196	4	8	8.5/9 = .944	.944	-1.59
235	2	4	4.5/5 = .900	(.944)(.900) = .850	-1.04
237	2	2	2.5/3 = .833	(.850)(.833) = .708	- .55

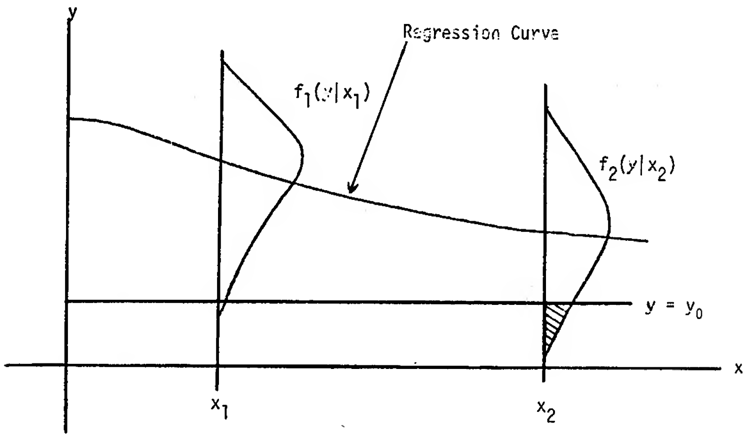


Figure 1. A Schematic Representation of the Aging Degradation Phenomenon

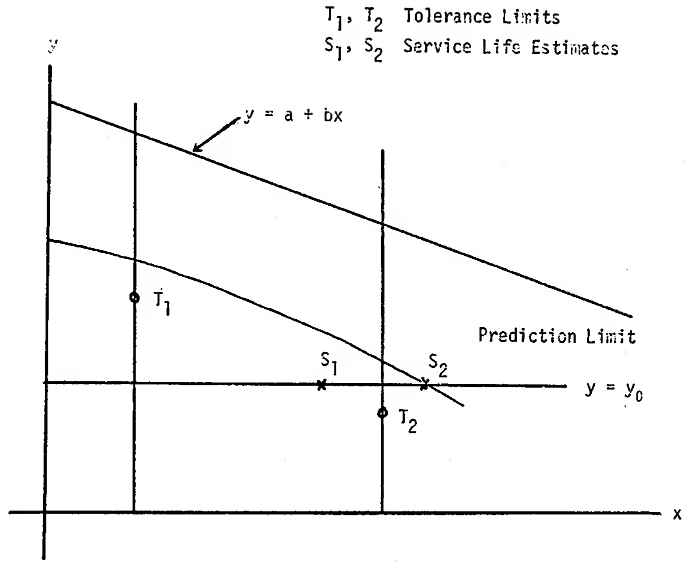
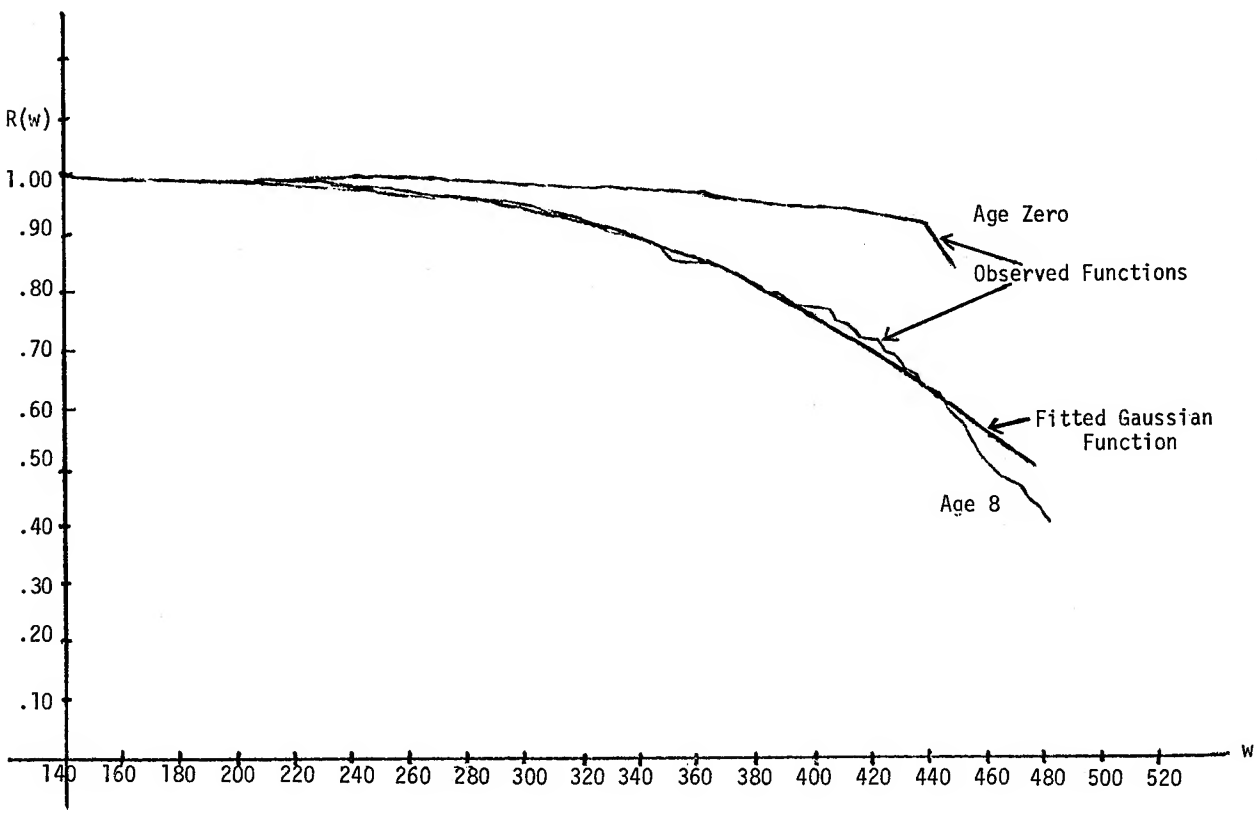


Figure 2. Hypothetical Service Life Estimates Obtained From Regression Analysis

Figure 3. Survival Functions



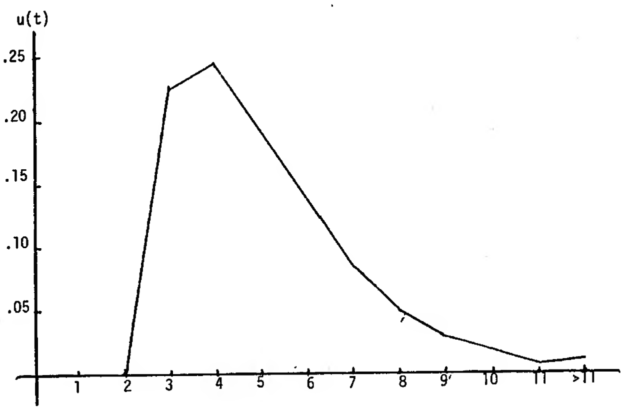


Figure 4. Time-To-Failure Density Function, u(t)

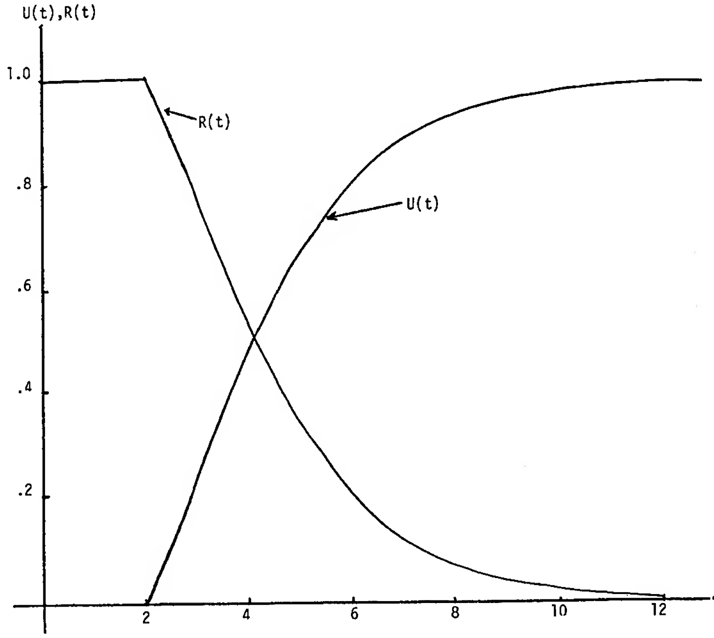


Figure 5. Reliability and Unreliability Functions

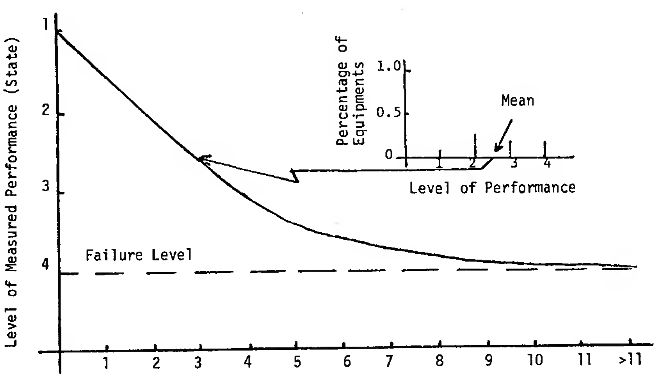


Figure 6. Average Deterioration Of A Performance Characteristic

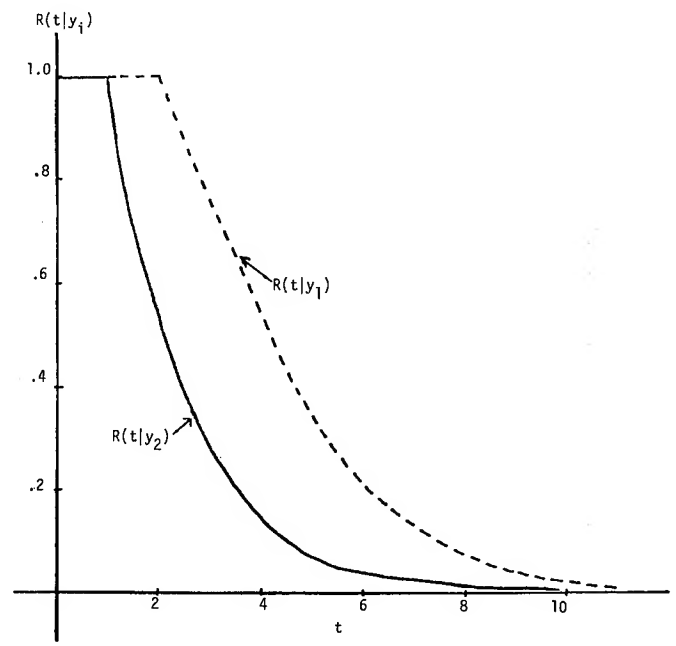


Figure 7. Conditional Reliability Functions

Thomas L. Paez
Instructor in Research
School of Civil Engineering
Purdue University
Lafayette, Indiana

Jhy-Pyng Tang
Assistant Professor of Civil Engineering
Purdue University
Lafayette, Indiana

James T. P. Yao
Professor of Civil Engineering
Purdue University
Lafayette, Indiana

Summary

A method is presented for computing the reliability of a structure with maintainable elements which may yield in response to strong excitations. Both general and specialized sequences of excitations are considered. The reliability of maintainable and non-maintainable structures is compared, and an expected cost criterion is also developed for use in comparison of the two types of structures.

Introduction

Most structural reliability studies have been formulated with the assumption that the condition of a structure can be described by two states, namely the safe state and the failure state (e.g., see[1,9]). This assumption does not allow differentiation between a structure which has experienced permanent damage but has not failed, and a structure which has not been damaged. This difference may be significant in some cases. It has been pointed out by Blume[2] that some structures might survive accelerations far in excess of code design requirements because of the energy absorption action by yielding and failure of minor parts of a structure. Moreover, Kasiraj and Yao[8] have shown that a structure can sustain a considerable amount of damage before failure finally occurs. Results of these studies indicate the importance of the consideration of a damage-state in the evaluation of structural reliability.

In this regard, Rosenblueth and Mendoza[3] developed a criterion for deciding whether or not a structure should be designed with elements which fail and absorb energy before the complete failure of the structure. Furthermore, several investigators[4,5,6,7] have studied the reliability of systems with standby redundancy. A solution to the reliability problem of a structure which accumulates damage may be possible using this concept, however, the idea of standby redundancy brings forward a more important concept. This is the concept of structural maintainability. If structures are assumed to accumulate damage, as they do in fact, what advantage can be gained by making a structure easily maintainable? This study is an attempt to explore in this direction for civil engineering structures.

In the following, the reliability of a maintainable structure is compared to that of a non-maintainable structure, where each is capable of accumulating permanent damage. The cumulative damage of each type of structure is expressed using a discrete representation, i.e., at any given time a structure might have i units of damage and once the cumulative damage in a structure exceeds L units, failure is assumed to occur. The number of units of damage in a structure will be called the damage level. In the following, the term structural damage will refer to damage which is not easily repairable; and the term maintainable damage, which is used with regard to the maintainable structure, will refer to damage which is easily repairable.

For illustration, an example of a maintainable structure which might actually be built is given. This is a typical steel frame with cross bracing as shown in Figure 1. The cross bracing might consist of cables or bars, which are designed in such a manner that they add some stiffness to the frame but will yield before the frame itself yields. These braces are also designed so that they can easily be replaced. Their main purpose is to absorb the energy of a strong motion excitation and lessen the possibility that the frame itself is damaged. A possible force displacement diagram for a maintainable structure is also given in Figure 1, along with a diagram of a non-maintainable structure and its typical force displacement diagram. In the following, the reliability of a structure is defined as the probability of survival when the structure is subjected to a sequence of strong motion excitations.

Reliability Analysis

To compare the maintainable structure with the non-maintainable type of structure, it is necessary to know (a) the types of excitation that these structures will be subjected to, (b) their responses to these excitations, and (c) the failure criterion for the structures. It is also desirable, in certain cases, to study the total cost of a maintainable structure versus that of a non-maintainable structure.

Although the types of dynamic excitation are varied and numerous, those types of excitations which have an important bearing on the analysis of reliability in a given structure can usually be identified. These significant excitations usually include earthquakes, wind loads, blast loads, etc. Assuming that (a) the significant types of excitations are specified and (b) the joint probability distribution for occurrence of these strong motion excitations along with some measure of their magnitudes and durations is specified, it is possible to specify the probability of occurrence of any sequence of excitations which describes a succession of various excitations in time. Let the probability of occurrence of a sequence of excitations be denoted by $P(S_i)$, $i = 1, 2, \ldots n_s$, then,

$$\sum_{i=1}^{n_s} P(S_i) = 1 \tag{1}$$

The concept of structural maintainability implies that there is advantage to repairing a damaged structure. As it has been pointed out previously, if damage repairs are easily performed, it might be advantageous to design structures so that some easily repairable damage occurs during extraordinarily strong excitations. The fact that maintainable structures derive their advantage from ease of damage repair necessitates their comparison to structures which can possibly sustain damage, but which cannot easily be repaired. And this in turn requires the use of a failure criterion which takes into consideration the possibility of failure due to cumulative damage.

In general, failure can occur due to one of the following two causes: (a) the strong motion excitation can cause some unacceptable level of permanent deformation in the structure; or (b) the cumulative damage can lead to fatigue failure. The reliability of a structure subjected to any sequence of strong-motion excitations in time can be computed if the

probability distribution of the structural response for each of the various types of excitation is known. In this study, failure due to the latter cause will be considered. Therefore, it is necessary to know (a) the probability that no damage occurs to the structure; (b) the probability distribution for the number of units of structural damage; and (c) the probability that only maintainable damage will occur when the structure under consideration is the maintainable type.

The reliability of a structure is found in the following manner. Define the damage transition probability matrix $[{}_jP]$, for the j^{th} excitation,

$$[{}_jP] = [{}_jp_{ab}] \qquad a, b = 0, 1, \dots L \tag{2}$$

The element ${}_jp_{ab}$ of the matrix $[{}_jP]$ denotes the probability that during the j^{th} excitation the amount of cumulative damage in the system changes from a units to b units. In this case, L denotes the maximum number of units of structural damage the system can withstand without failing.

Suppose that, a sequence of excitations $j_1, j_2, \dots j_n$ occurs, and denote this sequence by S_i. The damage transition matrix $[{}_{S_i}P]$ corresponding to the i^{th} sequence of excitations can be computed from the following formula.

$$[{}_{S_i}P] = [{}_{j_1}P][{}_{j_2}P]\dots[{}_{j_n}P] = \prod_{k=1}^{n} [{}_{j_k}P] \tag{3}$$

The element ${}_{S_i}p_{ab}$ of the matrix $[{}_{S_i}P]$ denotes the probability that the number of units of cumulative damage for the structure will change from a to b due to the sequence of excitations S_i.

A matrix $[P_o]$ is defined relating the probabilities that the structure begins in any given damage state. If $p_o(i)$ is the chance that the structure initially has i damage units, then $[P_o]$ is the diagonal matrix

$$[P_o] = [p_o(i)] \qquad i = 0, 1, \dots L \tag{4}$$

Premultiplying the matrix $[{}_{S_i}P]$ by the matrix $[P_o]$ we obtain

$$[{}_{oS_i}P] = [P_o][{}_{S_i}P] \tag{5}$$

The element ${}_{oS_i}p_{ab}$ of the matrix $[{}_{oS_i}P]$ denotes the probability that the structure will begin with a units of damage and end with b units due to the sequence of excitations S_i. The reliability of a structure subjected to the sequence of excitations S_i is the probability that no more than L units of damage have accumulated, which is the sum of the elements of $[{}_{oS_i}P]$.

$$R(S_i) = \sum_a \sum_b {}_{oS_i}p_{ab} \tag{6}$$

The overall reliability of the structure is then given by

$$R = \sum_i R(S_i)\, P(S_i) \tag{7}$$

where $P(S_i)$ has been previously defined as the probability of occurrence of the sequence of excitations S_i.

It is necessary to specify the damage transition probability matrices for the maintainable and non-maintainable structures.

For the maintainable structure, the probability that no damage occurs, ${}_jP_{om}$, is assumed to be known. Also, the probability that only repairable damage will occur, ${}_jp_D$, is assumed to be known. Finally, the probability distribution for the number of units of damage that occur is given by ${}_jp_m(i)$. Inclusion of the subscript j here denotes the association of these probabilities with the j^{th} excitation. The probabilities listed here can be used to infer the values of the elements of the damage transition matrix of Equation (2) for use in analysis of the reliability of a maintainable structure. These elements are obtained as follows: Assuming that negative damage cannot occur, the elements below the diagonal are zero.

$${}_jp_{ab} = 0 \qquad a>b \tag{8}$$

The chance that no structural damage occurs, or in other words, the chance that the damage state does not increase, is equal to the sum of the probability of no damage, p_o, and the probability of only repairable damage.

$${}_jp_{aa} = {}_jp_{om} + {}_jp_D \qquad a = 0, 1, \dots L \tag{9}$$

The chance that the quantity of damage sustained by the structures increases i units is equal to the probability that i units of damage occur.

$${}_jp_{ab} = {}_jp_m(b-a) \qquad \begin{array}{l} a = 0, 1, \dots L-1 \\ b = a + 1, \dots L \end{array} \tag{10}$$

Using the information in Equations (8), (9) and (10) the damage transition matrix for the maintainable structure subjected to the j^{th} excitation can be written as,

$$[{}_jP_m] = \begin{bmatrix} {}_jp_{om} + {}_jp_o & {}_jp_m(1) & {}_jp_m(2)\cdots {}_jp_m(L) \\ 0 & {}_jp_{om}+{}_jp_o & {}_jp_m(1) \quad \cdot \\ 0 & 0 & {}_jp_{om}+{}_jp_o \quad \cdot \\ \cdot & & \\ \cdot & & \\ 0 & \cdot \quad \cdot & \cdot \qquad {}_jp_{om}+{}_jp_o \end{bmatrix} \tag{11}$$

This can be used, along with the information on initial damage and excitation, to find the overall reliability of the maintainable structure.

For the non-maintainable structure, the probability that no damage occurs is assumed to be known, as is the probability distribution for the number of units of damage that occur. The former is denoted ${}_jp_{onm}$ and the latter ${}_jp_{nm}(i)$. The subscript j implies that these probabilities are valid for the j^{th} excitation. The elements of the damage transition probability matrix are obtained as follows. Since negative structural damage cannot occur, the below-diagonal elements are zero.

$${}_jp_{ab} = 0 \qquad a>b \tag{12}$$

The chance that there is no change in the amount of cumulative structural damage during the j^{th} excitation is the probability that no damage occurs, i.e.,

$${}_jp_{aa} = {}_jp_{onm} \qquad a = 0, 1, \dots L \tag{13}$$

The chance that the cumulative structural damage changes from a units to b units during the j^{th} excitation is the probability that b-a units of structural damage occur.

$$_jP_{ab} = {_jP_{nm}}(b-a) \qquad \begin{array}{l} a = 0, 1, \ldots L-1 \\ b = a + 1, \ldots L \end{array} \tag{14}$$

Equations (12), (13) and (14) yield the following damage transition probability matrix when used in Equation (2).

$$[{_jP_{nm}}] = \begin{bmatrix} {_jP_{onm}} & {_jP_{nm}}(1) & {_jP_{nm}}(2) \cdots & {_jP_{nm}}(L) \\ 0 & {_jP_{onm}} & {_jP_{nm}}(1) & \cdot \\ 0 & 0 & {_jP_{onm}} & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ 0 & \cdot & \cdot & {_jP_{onm}} \end{bmatrix} \tag{15}$$

Equation (15) pertains to the j^{th} excitation of a non-maintainable structure. This can be used, along with information about the initial damage and information about the excitation to compute the overall reliability of the non-maintainable structure.

Expressions for the overall reliabilities of the maintainable and non-maintainable types of structures have now been presented. If all the probabilities and probability distributions which have been used in this development are known, then the reliabilities of these two types of structural systems can be compared. One basis for comparison can be the expected total cost of each type of structure over the intended life span of the structure. Each must be subjected to the same excitations and have the same reliability. Let the initial cost of the non-maintainable structure be C_{nm}. Since no major structural maintenance is to be performed on the non-maintainable structure, C_{nm} is also the total cost over the total life span. Let the initial cost of the maintainable structure be denoted C_m. Since maintenance is performed on this structure every time damage occurs, there is a cost of maintenance αC_m. This is a fraction of the initial cost. Also, it is reasonable to assume the cost of maintenance being constant. So, if there are n damage - causing excitations in the life of the maintainable structure, the total cost of the structure is $C_m + \alpha n C_m = C_m (1 + \alpha n)$. However, the number of damage - causing excitations is a random variable. Therefore, the total cost of the maintainable structure is also a random variable. The average or expected cost of the maintainable structure can then be computed for the purposes of comparison.

In finding the probability distribution for the number of times that maintenance is performed on the structures, use will again be made of the structural response probabilities. As before, ${_jP_m}(i)$ is the probability that i units of structural damage occur without failure due to an excitation of the j^{th} type; ${_jP_D}$ is the probability of repairable damage; and ${_jP_{om}}$ is the probability that no damage whatsoever occurs. L is the maximum allowable number of units of structural damage. Define a matrix $[{_{jD}P}]$ as follows:

$$[{_{jD}P}] = [{_{jD}P_{ab}}], \quad a, b = 1, \ldots, L \tag{16}$$

The element ${_{jD}P_{ab}}$ of the above matrix represents the probability that some damage will occur (either structural damage or maintainable damage) due to the excitation j and that the level of structural damage will change from a to b. Since negative damage cannot occur

$${_{jD}P_{ab}} = 0 \text{ for } a>b \tag{17}$$

Since the occurrence of damage is required, only maintainable damage can occur if the damage state of the structure is to remain the same, so

$${_{jD}P_{aa}} = {_jP_D} \qquad a = 0, 1, \ldots L \tag{18}$$

For increases in the level of structural damage the elements of the matrix ${_{jD}P}$ are

$${_{jD}P_{ab}} = {_jP_m}\ (b-a) \qquad \begin{array}{l} a = 0, 1, \ldots L-1 \\ b = a + 1, \ldots L \end{array} \tag{19}$$

The matrix ${_{jD}P}$ can now be written

$$[{_{jD}P}] = \begin{bmatrix} {_jP_D} & {_jP_m}(1) & {_jP_m}(2) \cdots & {_jP_m}(L) \\ 0 & {_jP_D} & {_jP_m}(1) & \cdot \\ 0 & 0 & {_jP_D} & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ 0 & \cdot & \cdot & {_jP_D} \end{bmatrix} \tag{20}$$

Suppose that the structure is subjected to the sequence of excitations $j_1, j_2, \ldots j_n$. The probability that damage without failure occurs to the system k specific times out of the n total excitations is given by the sum of the elements in the matrix product of the k, $[{_{jD}P}]$'s for the k specific excitations during which damage is supposed to take place, multiplied by the product of the n-k, ${_jP_{om}}$'s corresponding to excitations during which no damage is supposed to take place. The probability that damage occurs <u>any</u> k times out of n is simply the sum of all the $\binom{n}{k}$ probabilities associated with a specific k damage occurrences. The formulas for this are written as follows. First write the probability, $P_o'(S_i)$, that damage does not occur on any of the excitations, given the sequence of excitations $j_1, j_2, \ldots j_n$. (S_i refers to this sequence).

$$P_o'(S_i) = {_{j_1}P_{om}}\ {_{j_2}P_{om}} \cdots {_{j_n}P_{om}} = \prod_{k=1}^{n} {_{j_k}P_{om}} \tag{21}$$

Then, the chance that no damage occurs, P_o, over all possible excitations is

$$P_o = \sum_i P_o'(S_i)\ P(S_i) \tag{22}$$

Now define a matrix which is a function of k arbitrary sequential excitations of the n total excitations

$$[P(j_1', \ldots j_k')] = [P_o]\ [{_{j_1'D}P_1}] \cdots [{_{j_k'D}P_1}] \tag{23}$$

The sum of the elements of this matrix can be denoted $P_S(j_1', \ldots, j_k')$ and is the chance that damage occurs during the k specific excitations $j'_1, \ldots, j_k'$, but that failure does not occur, and the probability of

no damage during the other excitations is

$$P''_o(S_i) = \prod_{j \neq j_1', \ldots, j_k'} {}_jp_{om} \tag{24}$$

So, the chance that damage occurs at exactly those k excitations corresponding to the j_i''s (j primes) and at none of the others is

$$P'(k, S_i) = P_S(j_1', \ldots, j_k') \cdot P''_o(S_i) \tag{25}$$

And the chance that damage occurs any k times in n excitations due to the sequence S_i of excitations is the sum of the $P'(k, S_i)$'s for all possible distinct sequential combinations that the $j_1', \ldots j_k'$ can take on.

$$P(k,S_i) = \sum_{\substack{\text{all combinations} \\ \text{of } j_i'\text{'s}}} P'(k,S_i) \tag{26}$$

The probability that k damages occur due to all possible sequences of excitations is

$$P_D(k) = \sum_i P(k,S_i)\, P(S_i) \tag{27}$$

Let N_m be the number of excitations in which damage occurs. Then the expected value of N_m is

$$E[N_m] = \sum_k k \cdot P_D(k) \tag{28}$$

Now a cost criterion for the effectiveness of the maintainable structure can be established. The cost of the non-maintainable structure is C_{nm}. The total expected cost of the maintainable structure is $C_m(1 + \alpha \cdot E[N_m])$. Let $C_m = \beta\, C_{nm}$. Then the maintainable structure is more desirable if and when.

$$C_m(1 + \alpha\, E[N_m]) < C_{nm}$$

or $$\beta(1 + \alpha\, E[N_m]) < 1 \tag{29}$$

If there exist α and β which satisfy this inequality, then the maintainable structure is to be chosen over the non-maintainable structure.

In all of the previous equations, a generalized set of excitations is considered. Frequently, one may be able to choose some single representative type of excitation in a conservative manner. In such a case, the equations are considerably simplified. In the following, consider only one type of excitation and let N_e be the number of occurrences of that excitation. Specify the probability distribution $P(N_e = k)$ for the member of occurrences of the excitation depending on the time under consideration.

Because all the excitations are assumed to be identical, the damage transition probability matrix as given by Equation (2) loses dependence on j and may be written

$$[P] = [p_{ab}] \qquad a, b = 0, 1, \ldots L \tag{30}$$

The element p_{ab} denotes the probability that for the assumed excitation, the damage state of the system changes from a to b; and L denotes the maximum number of units of structural damage the system can withstand without failing.

If a sequence of n excitations occurs, the n-step damage transition probability matrix is given by

$$[{}_nP] = [P]^n \tag{31}$$

The element ${}_np_{ab}$ of this matrix defines the probability that the system will accumulate b-a units of damage when subjected to n excitations. With the initial damage probability matrix defined as in Equation (4), a matrix $[{}_{on}P]$ can be defined.

$$[{}_{on}P] = [P_o][{}_nP] \tag{32}$$

The element ${}_{on}p_{ab}$ of the above matrix gives the probability that the system starts out with a units of damage and ends up with b units. The reliability of a structure subjected to n excitations is given by the sum of the elements of the matrix of Equation (32).

$$R(n) = \sum_a \sum_b {}_{on}p_{ab} \tag{33}$$

The overall reliability of the structure is given by

$$R = \sum_i R(i)\, P(N_e = i) \tag{34}$$

when the structure is subjected to only one type of excitation.

At this point, the damage transition matrices must be specified for the maintainable and non-maintainable structures.

For the maintainable structure, the development of the damage transition probability matrix is similar to that given in Equations (8) to (11). If the probability of no damage is p_{om}, the probability of maintainable damage p_D, and the probability distribution of structural damage $p_m(i)$, then the damage transition probability matrix for the maintainable structure is

$$[P_m] = \begin{bmatrix} p_{om} + p_D & p_m(1) & p_m(2)\cdots p_m(L) \\ 0 & p_{om} + p_D & p_m(1) & \cdot \\ 0 & 0 & p_{om} + p_D & \cdot \\ \cdot & & & \\ \cdot & & & \\ 0 & \cdot & \cdot & p_{om} + p_D \end{bmatrix} \tag{35}$$

This can be used in Equation (31) to find the reliability of a maintainable structure.

In the case of a non-maintainable structure, the development of the damage transition probability matrix is similar to that given in Equations (12) through (15). Let p_{onm} be the probability of no structural damage and let $p_{nm}(i)$ be the probability distribution of structural damage for the non-maintainable structures. Then the damage transition matrix is

$$[P_{nm}] = \begin{bmatrix} p_{onm} & p_{nm}(1) & p_{nm}(2)\cdots p_{nm}(L) \\ 0 & p_{onm} & p_{nm}(1) & \cdot \\ 0 & 0 & p_{onm} & \cdot \\ \cdot & & & \\ \cdot & & & \\ 0 & \cdot & \cdot & p_{onm} \end{bmatrix} \tag{36}$$

and when this is used in Equation (31), the reliability for a non-maintainable structure can be found.

The distribution of the number of times damage occurs is found in the following manner. Where there is only one type of excitation, define the matrix $[_D P]$.

$$[_D P] = [_D p_{ab}] \qquad a, b = 0, 1, \ldots L \tag{37}$$

The element $_D p_{ab}$ of the above matrix denotes the probability that damage occurs, due to the given excitation and that the damage states changes from a to b. If p_{om} is the probability that no damage occurs to the maintainable structure, p_D is the chance that only maintainable damage occurs, and $p_m(i)$ is the probability of occurrence of i units of structural damage, then $[_D P]$ is given by (see Equation (20)).

$$[_D P] = \begin{bmatrix} p_D & p_m(1) & p_m(2) \cdots & p_m(L) \\ 0 & p_D & p_m(1) & \cdot \\ 0 & 0 & p_D & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ 0 & \cdot & \cdot & p_D \end{bmatrix} \tag{38}$$

By raising the matrix of Equation (38) to the power k, the following result is obtained

$$[_{Dk} P] = [_D P]^k \tag{39}$$

The element $_{Dk} p_{ab}$ of the above matrix is the probability that damage occurs on k excitations and that the cumulative damage state changes from a to b. Premultiplying $[_{Dk} P]$ by $[P_o]$, we obtain,

$$[_{Dok} P] = [P_o][_{Dk} P] \tag{40}$$

The element $_{Dok} p_{ab}$ of this matrix is the chance that the system starts out with a units of structural damage, some type of damage occurs k times, and the system has b units of structural damage after the k^{th} time damage occurs. The sum of the elements of $[_{Dok} P]$ is the probability that the maintainable structure is damaged k times and survives. Let $\bar{P}(k)$ be the sum of the elements of $[_{Dok} P]$.

$$\bar{P}(k) = \sum_a \sum_b {}_{Dok} p_{ab} \tag{41}$$

The probability that the structure is not damaged is

$$\bar{P}_o = \sum_i p_{om}^{\,i} \; P(N_e = i) \tag{42}$$

Here, only one type of excitation has been considered and the distribution of its number of occurrences is given by $P(N_e = i)$.

The probability that the maintainable structure will be damaged k times is

$$_D\bar{P}(k) = \sum_{i=k}^{\infty} \binom{i}{k} \bar{P}(k) \; p_{om}^{\,i-k} P(N_e = i) \tag{43}$$

Here $P(N_e = i)$ is the probability if i excitations, $p_{om}^{\,i-k}$ is the chance that no damage occurs on i-k excitations and $\binom{i}{k}$ is the number of combinations of ways that damage can occur on k times out of i excitations.

The expected number of times that damage will occur is given by

$$E[N] = \sum_k k \cdot {}_D\bar{P}(k) \tag{44}$$

This can be used in the cost criterion for comparison of the maintainable and non-maintainable structures.

Numerical Examples

In the first example, the expected cost criterion is used for comparing a maintainable with a non-maintainable structure as shown in Figure 1. The required reliability for each structure is specified to be 0.9993, and the maximum number of damage units allowed is set equal to ten. The excitations acting on the structures are assumed to be of one type, following Possion arrivals with mean rate of occurrence being 0.2 yr^{-1}. And the time period of interest is specified as 10 years. Theoretically, there are an infinite number of designs which can be obtained with the desired reliability, however, a constraint was placed in the shape of the damage probability distributions in order to find one set of response probabilities each for the maintainable and the non-maintainable structures. The computed response probabilities are graphed in Figures 2 and 3 for the maintainable and non-maintainable structures respectively. Using the analysis presented in this paper, the expected number of times that damage occurs to the maintainable structure was found to be 1.197. Hence the criterion for choosing the maintainable structure over the non-maintainable is

$$\beta\,(1 + 1.197\,\alpha) < 1$$

$$\text{or} \quad \beta < (1 + 1.197\,\alpha)^{-1}$$

In other words, the initial cost of the maintainable structure must be less than $(1 + 1.197\,\alpha)^{-1}$ times the initial cost of the non-maintainable structure for the maintainable structure to be more advantageous. Figure 4 is a graph of β versus α for the given reliability value. It should be noted that the probability of no damage is 0.4 for the maintainable structure and 0.925 for the non-maintainable. It is hoped that this fact would easily enable the designer to satisfy the requirement that the initial cost of the non-maintainable structure be less than β times the initial cost of the maintainable structure. As a second example, though the maintainable and non-maintainable structures are designed with entirely different philosophies, an attempt is made to compare their reliabilities. Figure 5 is a graph showing the reliability of a maintainable structure after one to 20 excitations. The value of $p_{om} + p_D$ is noted on each curve and the distribution of damage was arbitrarily assumed to be linearly decreasing with a maximum of five units of damage. The damage at failure was taken to be 11 units. Figure 6 is a graph showing the reliability of a non-maintainable structure after one to twenty excitations. The value of p_{onm} is given corresponding to each curve, and the damage distribution was arbitrarily assumed to be linearly decreasing with a maximum of 10 units. Eleven units of damage were assumed to cause failure. Since the response probabilities do not change from excitation to excitation, the excitations were implicitly assumed to be identical.

Conclusion

A method has been presented for computing the reliability of a structure which is capable of responding to strong motion excitations with yielding. Formulas have been derived both for a very general type of loading and for a single type of loading with only random occurrences. Also, an expected cost criterion has been presented for use in comparing the maintainable and non-maintainable structures at a given reliability level.

It appears that, in at least some cases, the use of a maintainable type structure over a non-maintainable one can be highly advantageous. If the cost of repair for the maintainable structure is kept at a low enough level, then the maximum initial cost of the maintainable structure will be about the same as the initial cost of the non-maintainable structure. And, in general, the probability of no damage occurring will be much lower in the maintainable than in the non-maintainable structure, so, keeping its initial cost down should be an easy task for the designer.

The addition of maintainable members might also be used to increase the overall reliability of a structure, however, the effect of added bracing members for energy absorption has yet to be studied.

References

1. Shinozuka, M., "Methods of Safety and Reliability Analysis", Technical Report No. 1, Department of Civil Engineering and Engineering Mechanics, Columbia University, July 1969.

2. Blume, J.A., "Structural Dynamics in Earthquake Resistant Design", Journal of the Structural Division, ASCE, Vol. 84, No. ST4, Part 1, Proc. Paper 1695, July 1958, pp. 1-45.

3. Rosenblueth, M., and Mendoza, E., "Reliability Optimization in Isostatic Structures", Journal of the Mechanics Division, ASCE, Vol. 97, No. EM6, December, 1971.

4. Gnedenko, B.V., et. al., Mathematical Methods of Reliability Theory, Academic Press, New York, 1969.

5. Barlow, R.E., and Proschan, F., Mathematical Theory of Reliability, John Wiley and Sons, Inc., New York, 1965.

6. Calabro , S.R., Reliability Principles and Practices, McGraw-Hill Book Co., New York, 1962.

7. Brazovsky, I., Reliability Theory and Practice, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1961.

8. Kasiraj, I., and Yao, J.T.P., "Fatigue Damage in Seismic Structures", Journal of the Structural Division, ASCE, Vol. 95, No. ST8, August 1969.

9. Ang, A.H.-S., and Amin, M., "Reliability of Structures and Structural Systems", Journal of the Engineering Mechanics Division, ASCE, Vol. 94, No. EM2, Proc. Paper 5919, April 1968, pp. 671-691.

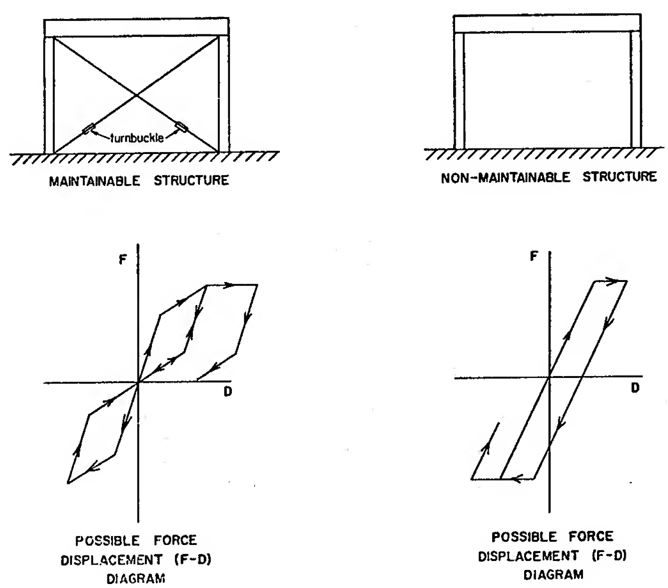
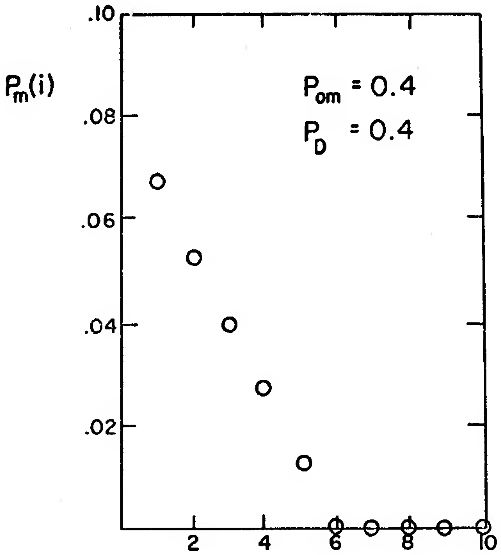
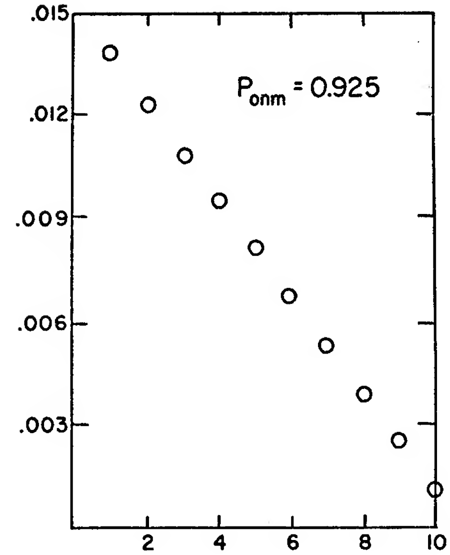


FIGURE I.



RESPONSE PROBABILITIES
FOR THE MAINTAINABLE
STRUCTURE

FIGURE 2.



RESPONSE PROBABILITIES
FOR THE NON-MAINTAINABLE
STRUCTURE

FIGURE 3.

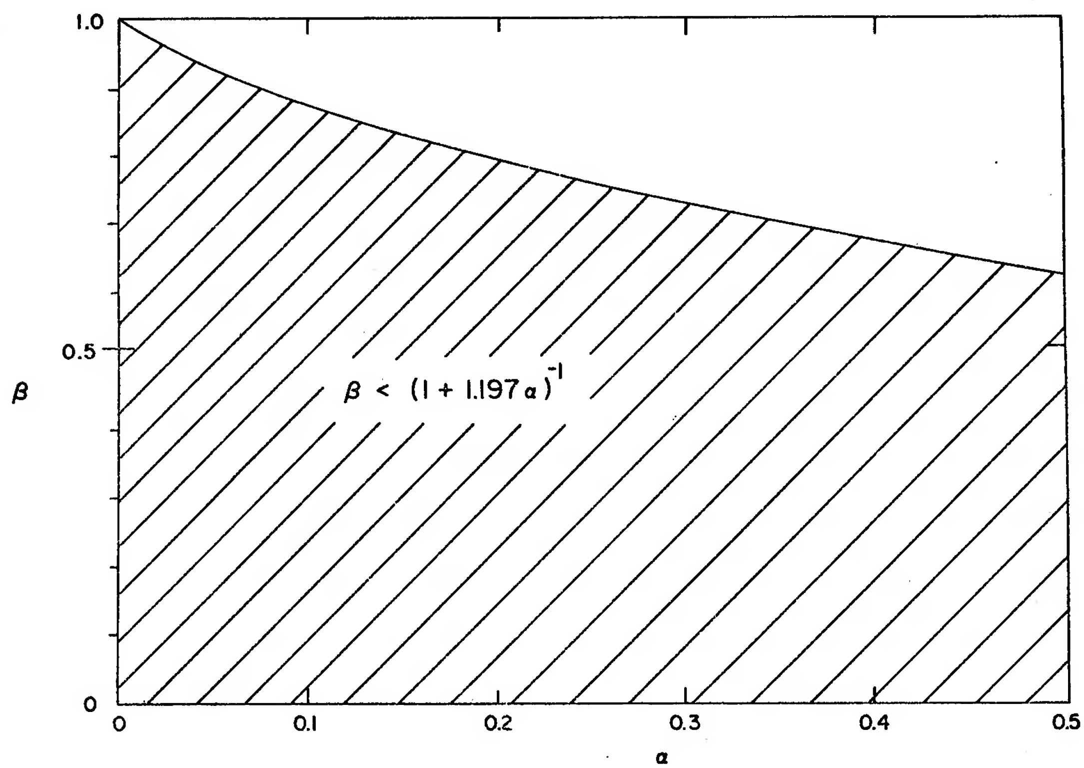


FIGURE 4. EXPECTED COST CRITERION FOR R = 0.9993

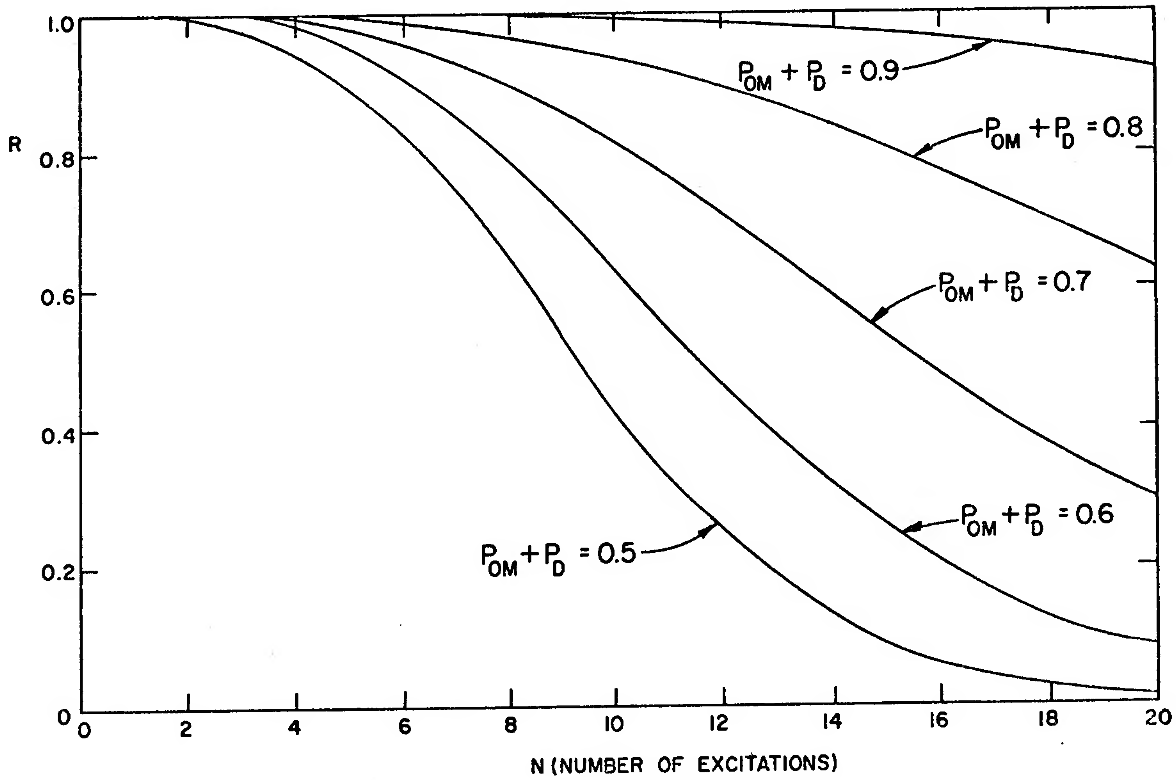


FIGURE 5. RELIABILITY OF A MAINTAINABLE STRUCTURE

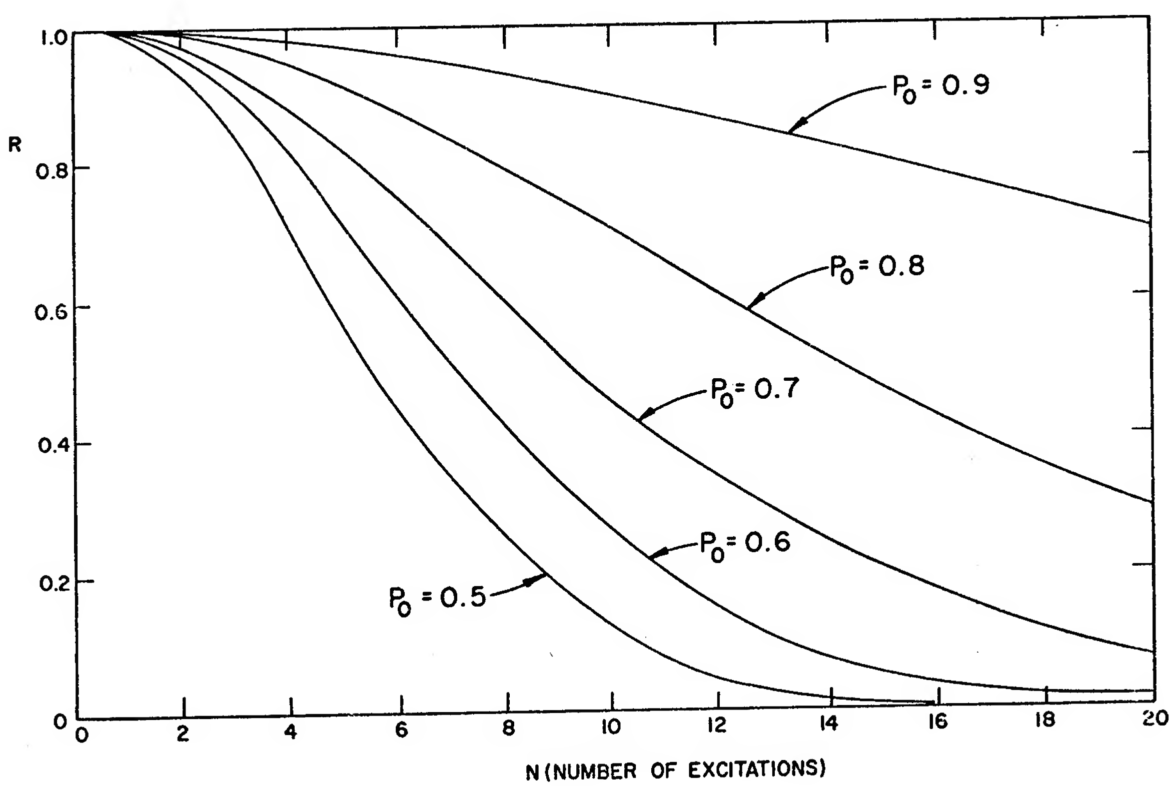


FIGURE 6. RELIABILITY OF NON-MAINTAINABLE STRUCTURE

SOME APPLICATIONS OF INTEGRAL TRANSFORMS IN RELIABILITY*

RAM D. PRASAD
IBM Corporation
Armonk, New York 10504

Introduction

Integral transform theory has been used to derive the compound binomial distribution when the parameter p (the probability of success) is the product of n independent random variables p_i's having any general distributions. Simplified expressions and the asymptotic representation for the probability distribution function have been obtained for specific distributions of p_i's. The equivalence relationship between binomial and Pascal distributions has been stated and proved.

Applications of these results are shown in reliability of series-parallel systems when the probability of success of each component is independent random variable and in determination of quantity to start in an n stage manufacturing process having random operation yields in order to obtain a specified quantity of finished goods.

1. Compound Binomial Distribution

Consider the binomial random variable X, given by the probability function,

$$Pr[X=k] = \binom{n}{k} p^k (1-p)^{n-k}, \quad k=1,2,\cdots n,$$

where k is number of success in n trials, and p is the probability of achieving success in a single trial.

The probability distribution function of X is given by

$$Pr[X \le c] = \sum_{k=0}^{c} \binom{n}{k} p^k (1-p)^{n-k}.$$

Let p be the product of N random variables $p_1, p_2, \ldots, p_N$ defined in the interval $(a_i, b_i) \subseteq (0,1)$ then the conditional distribution of X is given by

$$Pr[X \le c/p] = \sum_{k=0}^{c} \binom{n}{k} p^k (1-p)^{n-k} \qquad \ldots(1.1)$$

Consider the well known identity

$$\sum_{k=0}^{c} \binom{n}{k} p^k (1-p)^{n-k} = (n-c)\binom{n}{c} \int_p^1 y^c (1-y)^{n-c-1} dy. \qquad \ldots(1.2)$$

*The results reported in this paper are part of the author's doctoral dissertation, "Some Results on Integral Transforms and Their Application to Certain Reliability and Stochastic Linear Programming Problems," submitted to the Department of Industrial Engineering and Operations Research, New York University, May, 1972.

From (1.1) and (1.2)

$$Pr[X \le c/p] = (n-c)\binom{n}{c} \int_p^1 y^c (1-y)^{n-c-1} dy,$$

$$Pr[X \le c] = \int_0^1 Pr[X \le c/p]\, f(p)\, dp$$

$$= K \int_0^1 f(p)\, dp \int_p^1 y^c (1-y)^{n-c-1} dy,$$

$$K = (n-c)\binom{n}{c}$$

$$= \lim_{\alpha \to 1} K \int_0^1 p^{\alpha-1} f(p)\, dp \int_p^1 y^c (1-y)^{n-c-1} dy$$

$$= \lim_{\alpha \to 1} \frac{K}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} d\omega\, M_p(\alpha - \omega) \cdot \frac{\Gamma(n-c)\,\Gamma(\omega+c+1)}{\omega\,\Gamma(\omega+n+1)},$$

where $M_p(\alpha)$ is the Mellin transform of the random variable p, the Mellin transform of

$$\int_0^1 y^c (1-y)^{n-c-1} dy = \frac{\Gamma(n-c)\,\Gamma(\alpha+c+1)}{\alpha\,\Gamma(\alpha+n+1)}.$$

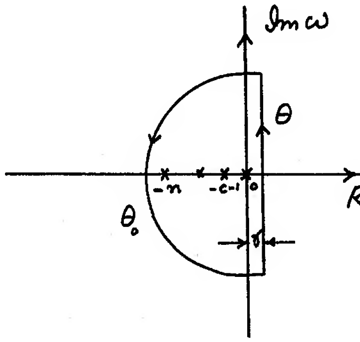
Therefore,

$$Pr[X \le c] = \lim_{\alpha \to 1} \frac{1}{2\pi i} \frac{\Gamma(n+1)}{\Gamma(c+1)} \cdot \int_{\gamma - i\infty}^{\gamma + i\infty} \frac{M_p(\alpha-\omega)}{\omega} \frac{\Gamma(\omega+c+1)}{\Gamma(\omega+n+1)} d\omega$$

$$= \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} \frac{M_p(1-\omega)}{\omega} \prod_{k=1}^{n-c} \left(\frac{c+k}{\omega+c+k}\right) d\omega. \qquad \ldots(1.3)$$

Eq. (1.3) is the integral representation of compound binomial distribution.

$M_p(1-\omega)$ is analytic in $Re\,\omega \le 0$ and $|M_p(1-\omega)| \le Q$, $Q > 0$, for $Re\,\omega \le 0$, since $M_p(\omega)$ is the Mellin transform of the probability density function in the interval $\le (0,1)$. Thus the complex integral in (1.3) can be evaluated on the contour shown in Fig. 1.

Fig.1 

Consider the closed contour θ , consisting of the straight line parallel to the imaginary axis and at a distance γ to the left of it and the semicircle θ_0 with center (γ, o) , and the radius R , γ has been chosen in such a way that all the singularities of the integrand lie to the left of the line joining $\gamma - i\infty$ to $\gamma + i\infty$.

On θ_0 ; $\omega = R\exp(i\theta)$ and

$$\left| \frac{M_p(1-\omega)}{\omega} \prod_{k=1}^{n-c} \frac{c+k}{\omega+c+k} \right| \le \frac{Q}{|\omega|} \prod_{k=1}^{n-c} \frac{c+k}{|\omega+c+k|},$$

$$\le \frac{Q}{R} \prod_{k=1}^{n-c} \frac{c+k}{[(R\cos\theta+c+k)^2 + R^2\sin^2\theta]^{1/2}}$$

$\to 0$ as $R \to \infty$, and therefore,

$$\lim_{R\to\infty} \left| \int_{\theta_0} \frac{M_p(1-\omega)}{\omega} \prod_{k=1}^{n-c} \frac{c+k}{\omega+c+k}\, d\omega \right| = 0,$$

and the integral (1.3) can be written as

$$Pr[X \le c] = \frac{1}{2\pi i} \oint_{\theta} \frac{M_p(1-\omega)}{\omega} \prod_{k=1}^{n-c} \left(\frac{c+k}{\omega+c+k}\right) d\omega.$$

...(1.4)

The integrand in (1.4) is analytic on and inside θ , except at finite number of isolated points on $-ve$ real axis (i.e., at $\omega = 0$, $-c-1$, $-c-2$, ..., $-n$), and therefore by Cauchy-residue Theorem (1.4) can be written as

$$Pr[X \le c] = 1 - \sum_{r=1}^{n-c} M_p(1+c+r) \prod_{k=1,\ne r}^{n-c} \left(\frac{k+c}{k-r}\right),$$

...(1.5)

$$\text{or } Pr[X > c] = \sum_{r=1}^{n-c} M_p(1+c+r) \prod_{k=1,\ne r}^{n-c} \left(\frac{k+c}{k-r}\right).$$

...(1.6)

Let $f(p_1, p_2, \dots, p_N)$ be joint density function of random variables (p_1 , p_2 , ..., p_N). Then,

$$M_p(\alpha) = \iint_{(N)} \cdots \int \left(\prod_{i=1}^{N} p_i\right)^{\alpha-1} f(p_1, p_2, \dots p_N)$$

$$dp_1 . dp_2 \cdots dp_N \ . \quad ...(1.7)$$

If $p = p_1 \cdot p_2 \cdots p_N$, and p_1, p_2, ..., p_N are statistically independent, then

$$M_p(\alpha) = \prod_{i=1}^{N} M_{p_i}(\alpha) \ ,$$

...(1.8)

where $M_{p_i}(\alpha)$ is the Mellin transform of the random variable p_i . And the compound binomial distribution function is given by

$$Pr[X \le c] = 1 - \sum_{r=1}^{n-c} \prod_{i=1}^{N} M_{p_i}(1+c+r) \prod_{k=1,\ne r}^{n-c} \left(\frac{k+c}{k-r}\right).$$

...(1.9)

Expression (1.9) can be further simplified as follows:

$$\prod_{k=1,\ne r}^{n-c} \left(\frac{k+c}{k-r}\right)$$

$$= \frac{(1+c)(2+c)\cdots\cdots(r+c-1)(r+c+1)\cdots\cdots n}{(1-r)(2-r)\cdots\cdots(-2)(-1)\ 1.2.\cdots\cdots(n-r-c)}$$

$$= \frac{(-c-1)(-c-2)\cdots\cdots(-r-c+1)}{1.2.3\cdots\cdots(r-1)} \ \frac{n!}{(r+c)!\,(n-r-c)!}$$

$$= \binom{-c-1}{r-1}\binom{n}{r+c}. \quad ...(1.10)$$

From (1.9) and (1.10),

$$Pr[X \le c]$$

$$= 1 - \sum_{r=1}^{n-c} M_p(1+c+r)\binom{-c-1}{r-1}\binom{n}{r+c}$$

$$= 1 - \sum_{r=1}^{n-c} (-1)^{r-1}\binom{r+c-1}{r-1}\binom{n}{r+c} M_p(1+c+r)$$

$$= 1 - \sum_{r=0}^{n-c-1} (-1)^{r}\binom{r+c}{r}\binom{n}{r+c+1} M_p(2+c+r)$$

$$= 1 - \sum_{r=0}^{n-c-1} (-1)^{r} \frac{(r+c)!}{r!\,c!} \ \frac{n!\,M_p(2+c+r)}{(r+c+1)!\,(n-r-c-1)!},$$

$$Pr\,[X^{\circ} \leq c]$$

$$= 1 - \frac{n!}{c!} \sum_{r=0}^{n-c-1} \frac{(-1)^r}{r!\,(n-r-c-1)!} \frac{M_p(2+c+r)}{1+c+r}$$

$$= 1 - \frac{n!}{c!\,(n-c-1)!} \sum_{r=0}^{n-c-1} (-1)^r \binom{n-c-1}{r} \frac{M_p(2+c+r)}{1+c+r}$$

$$= 1 - K \sum_{r=0}^{n-c-1} (-1)^r \binom{n-c-1}{r} \frac{\prod_{i=1}^{N} M_{p_i}(2+c+r)}{1+c+r},$$

where $K = \dfrac{n!}{c!\,(n-c-1)!} = \dfrac{1}{B(c+1, n-c)}$ ...(1.11)

Eq. (1.11) represents the compound binomial distribution when the parameter p is the product of N independent random variables.

Approximations to Compound Binomial Distribution

2. Binomial-Beta Distribution

The binomial-beta distribution is a compound binomial distribution when the parameter p is a random variable having beta distribution. Its asymptotic behavior was obtained by Dubey [3] in connection with compound Pascal distribution. It was shown in [3] that the binomial-beta distribution can be approximated by the binomial distribution with $p = a/(a+b)$, provided the parameters a and b of the beta distribution are large. Similar limit theorem has been proved in this section when the parameter p (of the binomial distribution) is the product of N independent random variables p_j's $(j = 1, 2, \ldots N)$ have beta distribution.

Let the p.d.f. of p_j be given by

$$f_{p_j}(t) = \frac{t^{a_j-1}\,(1-t)^{b_j-1}}{B(a_j, b_j)}, \quad 0<t<1. \qquad \text{...(2.1)}$$

The Mellin transform of p_j is well known and given by

$$M_{p_j}(\alpha) = \frac{B(b_j, a_j+\alpha-1)}{B(a_j, b_j)}. \qquad \text{...(2.2)}$$

Hence

$$M_p(\alpha) = \prod_{j=1}^{N} \frac{B(b_j, a_j+\alpha-1)}{B(a_j, b_j)}$$

$$= \prod_{j=1}^{N} \left[\frac{\Gamma(a_j+\alpha-1)}{\Gamma(a_j)} \frac{\Gamma(a_j+b_j)}{\Gamma(a_j+b_j+\alpha-1)}\right]. \qquad \text{...(2.3)}$$

Using the well known asymptotic result, namely,

$$\Gamma(x+a)/\Gamma(x) \sim x^a, \quad x \to \infty,$$

equation (2.3) can be written as

$$M_p(\alpha) = \prod_{j=1}^{N} \left(\frac{a_j}{a_j+b_j}\right)^{\alpha-1}, \ \{a_j, b_j\} \to \infty. \qquad \text{...(2.4)}$$

From (1.11), the compound binomial distribution is given by

$$Pr\,[X \leq c] = 1 - K \sum_{r=0}^{n-c-1} \frac{(-1)^r}{r+c+1} \binom{n-c-1}{r} M_p(2+c+r)$$

$$\sim 1 - K \sum_{r=0}^{n-c-1} \frac{(-1)^r}{r+c+1} \binom{n-c-1}{r} \prod_{j=1}^{N} \left(\frac{a_j}{a_j+b_j}\right)^{r+c+1}.$$

Let $\omega = \prod_{j=1}^{N} \left(a_j/(a_j+b_j)\right)$, then

$$Pr\,[X \leq c] \sim 1 - K \sum_{r=0}^{n-c-1} \frac{(-1)^r}{r+c+1} \binom{n-c-1}{r} \omega^{r+c+1}$$

$$= 1 - K \int_0^{\omega} \sum_{r=0}^{n-c-1} (-1)^r \binom{n-c-1}{r} x^{r+c}\,dx$$

$$= 1 - K \int_0^{\omega} x^c (1-x)^{n-c-1}\,dx$$

$$= K \int_{\omega}^{1} x^c (1-x)^{n-c-1}\,dx. \qquad \text{...(2.5)}$$

Clearly the limiting form of the compound binomial distribution is binomial with parameter $p = \prod_{j=1}^{N} (a_j/(a_j+b_j))$ when $\{a_j, b_j\}$ are large.

In the special case, when $N=1$, the result obtained in this section compares with the result for binomial-beta distribution derived in [3]. Thus the following limit theorem can be stated for the compound binomial distribution.

Theorem (2.1) The compound binomial distribution when the parameter p is the product of N independent random variables p_j's, $(j = 1, 2, \ldots N)$ having beta distributions with parameters $\{a_j, b_j\}$ is approximately binomial provided $\{a_j, b_j\}$ are large.

3. Binomial-Uniform Distribution

The compound binomial distribution when parameter p is product of N independent and identical random variables having uniform distribution on the unit support will be derived in this section.

Let the p.d.f. of p_j be given by

$$f_{p_j}(p_j) = 1, \quad 0 \le p_j \le 1, \quad j = 1, 2, \cdots N. \qquad \ldots(3.1)$$

The Mellin transform of p_j is given by

$M_{p_j}(\alpha) = \frac{1}{\alpha}$ and therefore

$$M_p(\alpha) = \frac{1}{\alpha^N}. \qquad \ldots(3.2)$$

From (1.11), the compound binomial distribution is given by

$$Pr[X \le c] = 1 - K \sum_{r=0}^{n-c-1} \frac{(-1)^r}{r+c+1} \binom{n-c-1}{r} M_p(2+c+r)$$

$$= 1 - K \sum_{r=0}^{n-c-1} (-1)^r \binom{n-c-1}{r} \frac{1}{(r+c+1)(r+c+2)^N}$$

$$= 1 - K \sum_{r=0}^{n-c-1} (-1)^r \binom{n-c-1}{r} \int_0^\infty \exp(-(r+c+1)t) \frac{\gamma(N,t)}{\Gamma(N)} dt,$$

where $\gamma(N,t) = \int_0^t e^{-x} x^{N-1} dx$, or

$$Pr[X \le c] = 1 - \frac{K}{\Gamma(N)} \int_0^\infty \gamma(N,t) e^{-(c+1)t} (1-e^{-t})^{n-c-1} dt. \qquad \ldots(3.3)$$

Consider the following Laplace transforms

$$\int_0^\infty e^{-st} \gamma(N,t)\, dt = \frac{\Gamma(N)}{s(s+1)^N}$$

$$\int_0^\infty e^{-st} (1-e^{-t})^{n-c-1} dt = B(s, n-c).$$

Using the property of the Laplace transform of product functions, (3.3) can be written as

$$Pr[X \le c] = 1 - \frac{K}{2\pi i} \int_{-i\infty}^{i\infty} \frac{B(c+1-s, n-c)}{s(s+1)^N} ds. \qquad \ldots(3.4)$$

In $Re(s) \le 0$, the singularities of the integrand of (3.4) are a simple pole at $s=0$ and a multiple pole of multiplicity N at $s=-1$; therefore, by residue theorem, (3.4) reduces to

$$Pr[X \le c] = 1 - K\Big[B(c+1, n-c) + \frac{1}{\Gamma(N)} \frac{d^{N-1}}{ds^{N-1}} \Big(\frac{1}{s} B(c+1-s, n-c)\Big)_{s=-1}\Big]$$

$$= \frac{K}{\Gamma(N)} \frac{d^{N-1}}{ds^{N-1}} \Big[\frac{1}{s} B(c+1-s, n-c)\Big]_{s=-1}$$

$$= \frac{K}{\Gamma(N)} \sum_{k=0}^{N-1} \binom{N-1}{k} B^{(k)}(c+2, n-c) (-1)^k (N-k-1)!$$

$$= \frac{1}{B(c+1, n-c)(N-1)!} \sum_{k=0}^{N-1} \frac{(N-1)!\,(-1)^k}{k!} B^{(k)}(c+2, n-c)$$

$$= \frac{1}{B(c+1, n-c)} \sum_{k=0}^{N-1} \frac{(-1)^k}{k!} B^{(k)}(c+2, n-c). \qquad \ldots(3.5)$$

The asymptotic behavior of the binomial uniform distribution for large values of n and c can be obtained as follows:

From (3.4), the binomial-uniform distribution is given by

$$Pr[X \le c] = 1 - \frac{K}{2\pi i} \int_{-i\infty}^{i\infty} \frac{ds}{s(s+1)^N} B(c+1-s, n-c)$$

$$= 1 - \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{ds}{s(s+1)^N} \frac{\Gamma(c+1-s)\Gamma(n+1)}{\Gamma(c+1)\Gamma(n+1-s)}$$

$$\sim 1 - \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{ds}{s(s+1)^N} \Big(\frac{c+1}{n+1}\Big)^{-s},$$

$$(n, c \to \infty).$$

The inverse of the Mellin transform $(s(s+1)^N)^{-1}$ is

$$\gamma(N, -\ln x)/\Gamma(N), \quad 0 < x < 1,$$

and therefore

$$Pr[X \le c] \sim 1 - \gamma(N, -\ln(\tfrac{c+1}{n+1}))/\Gamma(N)$$

$$= 1 - \gamma(N, \ln(\tfrac{n+1}{c+1}))/\Gamma(N)$$

$$= \Gamma(N, \ln(\tfrac{n+1}{c+1}))/\Gamma(N). \qquad \ldots(3.6)$$

where $\Gamma(a,x)$ is incomplete gamma function and is defined by

$$\Gamma(a,x) = \int_x^\infty e^{-t}\, t^{a-1} dt .$$

Thus the following limit theorem:

Theorem (3.1) The compound binomial distribution when the parameter p is the product of N independent and identical random variables having uniform distribution on unit support is asymptotically given by

$$Pr[X \le c] \sim \Gamma(N, \ln(\tfrac{n+1}{c+1}))/\Gamma(N)$$

as $(c,n) \to \infty$.

As a special case, consider $N=1$. Eq. (3.5) yields

$$Pr[X \le c] = \frac{B(c+2, n-c)}{B(c+1, n-c)} = \frac{c+1}{n+1} .$$

Also

$$Pr[X=c] = \frac{1}{n+1} . \qquad \ldots(3.7)$$

From the asymptotic result

$$Pr[X \le c] \sim \Gamma(1, \ln(\tfrac{n+1}{c+1}))$$

$$= \exp[-\ln(\tfrac{n+1}{c+1})] = \frac{c+1}{n+1} .$$

Thus the result (3.7) agrees with the well known result obtained by direct integration.

Next consider the case $N=2$. From (3.5) the binomial-uniform distribution is given by

$$Pr[X \le c] = K[B(c+2, n-c) - B^{(1)}(c+2, n-c)]$$

$$= K\Big[B(c+2, n-c) - B(c+2, n-c)\{\psi(c+2) - \psi(n+2)\}\Big]$$

$$= K\, B(c+2, n-c)[1 + \psi(n+2) - \psi(c+2)]$$

$$= \frac{c+1}{n+1}\Big[1 + \sum_{k=c+2}^{n+1} \frac{1}{k}\Big] . \qquad \ldots(3.8)$$

And

$$Pr[X=c] = Pr[X \le c] - Pr[X \le c-1]$$

$$= \frac{c+1}{n+1}\Big[1 + \sum_{k=c+2}^{n+1}\frac{1}{k}\Big] - \frac{c}{n+1}\Big[1 + \sum_{k=c+1}^{n+1}\frac{1}{k}\Big]$$

$$= \frac{c+1}{n+1}\Big[1 + \sum_{k=c+1}^{n+1}\frac{1}{k} - \frac{1}{c+1}\Big] - \frac{c}{n+1}\Big[1 + \sum_{k=c+1}^{n+1}\frac{1}{k}\Big]$$

$$= \frac{1}{n+1}\sum_{k=c+1}^{n+1}\frac{1}{k}$$

$$= \frac{\psi(n+2) - \psi(c+1)}{n+1} \qquad \ldots(3.9)$$

From asymptotic formula (3.6)

$$Pr[X \le c] \sim \frac{\Gamma(2, \ln(\tfrac{n+1}{c+1}))}{\Gamma(2)}$$

$$= \exp[-\ln(\tfrac{n+1}{c+1})][1 + \ln(\tfrac{n+1}{c+1})]$$

$$= \frac{c+1}{n+1}[1 + \ln(\tfrac{n+1}{c+1})] \qquad \ldots(3.10)$$

When $N=3$, Eq. (3.5) yields

$$Pr[X \le c] = K\Big[B(c+2, n-c) - B^{(1)}(c+2, n-c) + \tfrac{1}{2}B^{(2)}(c+2, n-c)\Big]$$

$$= K\, B(c+2, n-c)\Big[1 + \psi(n+2) - \psi(c+2) + \tfrac{1}{2}\{(\psi(c+2) - \psi(n+2))^2 + \psi'(c+2) - \psi'(n+2)\}\Big] . \qquad \ldots(3.11)$$

Making use of the property of polygamma function (ref [8]), namely

$$\psi'(k+1) = \zeta(2) - \sum_{j=1}^{k} (j)^{-2}$$

, where $\zeta(x)$ is Rieman Zeta function, (3.11) reduces to

$$Pr[X \le c] = \frac{c+1}{n+1}\left[1 + \sum_{k=c+2}^{n+1} \frac{1}{k} + \frac{1}{2}\left\{\left(\sum_{k=c+2}^{n+1} \frac{1}{k}\right)^2 + \sum_{k=c+2}^{n+1} \frac{1}{k^2}\right\}\right], \quad \text{...(3.12)}$$

and

$$Pr[X = c] = \frac{1}{2(n+1)}\left[\left(\sum_{k=c+1}^{n+1} \frac{1}{k}\right)^2 + \sum_{k=c+1}^{n+1} \frac{1}{k^2}\right]$$

$$= \frac{1}{2(n+1)}\left[\{\psi(c+1) - \psi(n+2)\}^2 + \psi'(c+1) - \psi'(n+2)\right]. \quad \text{...(3.13)}$$

From asymptotic formula (3.6),

$$Pr[X \le c] \sim \frac{\Gamma(3, \ln(\frac{n+1}{c+1}))}{\Gamma(3)} \quad \text{...(3.14)}$$

$$= \exp\left[-\ln\left(\frac{n+1}{c+1}\right)\right]\cdot\left(1 + \ln\left(\frac{n+1}{c+1}\right) + \frac{1}{2}\left(\ln\left(\frac{n+1}{c+1}\right)\right)^2\right)$$

$$= \frac{c+1}{n+1}\left[1 + \ln\left(\frac{n+1}{c+1}\right) + \frac{1}{2}\left(\ln\left(\frac{n+1}{c+1}\right)\right)^2\right]. \quad \text{...(3.15)}$$

4. The Binomial-Gamma Distribution

The binomial-gamma distribution has been defined by Dubey [3], in connection with the Pascal-Gamma distribution. The conditional p.f. $f(X/\omega)$ of the binomial random variable X was represented by

$$f(X/\omega) = Pr[X = c/\omega] = \binom{n}{c}(e^{-\omega})^c (1 - e^{-\omega})^{n-c}, \quad c = 0, 1, \ldots n,$$

where ω is the random variable whose p.d.f. is given by

$$f_\Omega(\omega) = \frac{\lambda^\beta \omega^{\beta-1} e^{-\lambda\omega}}{\Gamma(\beta)}, \quad \omega > 0, \ (\lambda, \beta > 0).$$

The above representation can be generalized as follows:

Let $p_i = \exp(-\omega_i)$, and p be a function of random variables p_j's, $(j = 1, 2, \ldots N)$. For the purpose of illustration let $\{\omega_j\}$ are independent and $p = p_1 \cdot p_2 \cdot \ldots \cdot p_N$.

The conditional distribution function o of the binomial variate X is now represented by

$$Pr[X \le c / \omega_1, \omega_2, \ldots, \omega_N] = \sum_{k=0}^{c} \binom{n}{k} p^k (1-p)^{n-k} = K \int_p^1 y^c (1-y)^{n-c-1} dy. \quad \text{...(4.1)}$$

and the p.d.f. of ω_j is given by

$$f_{\Omega_j}(\omega) = \lambda_j^{\beta_j} \omega^{\beta_j - 1} \exp(-\lambda_j \omega)/\Gamma(\beta_j), \quad \omega > 0,$$

for $j = 1, 2, \ldots, N$.

The Mellin transform of P_j is given by

$$M_{P_j}(\alpha) = \int_0^1 p_j^{\alpha-1} f_{P_j}(p_j)\, dp_j = \int_0^\infty e^{-(\alpha-1)\omega_j} f_{\Omega_j}(\omega_j)\, d\omega_j = L_{\Omega_j}(\alpha - 1). \quad \text{...(4.2)}$$

where $L_{\Omega_j}(s)$ is the Laplace transform of ω_j.

Hence

$$M_{P_j}(\alpha) = \left(\frac{\lambda_j}{\lambda_j + \alpha - 1}\right)^{\beta_j} \quad \text{...(4.3)}$$

From (1.11), the binomial-gamma distribution is thus given by

$$Pr[X \le c] = 1 - K \sum_{r=0}^{n-c-1} \frac{(-1)^r}{r+c+1} \cdot \binom{n-c-1}{r} \prod_{j=1}^{N} \left(\frac{\lambda_j}{\lambda_j + r + c + 1}\right)^{\beta_j}. \quad \text{...(4.4)}$$

Consider a special case when $\{\omega_j\}$ are identical random variables with $\{\lambda_j\} = \lambda$ and $\{\beta_j\} = \beta$. Then the binomial-gamma distribution is given by

$$Pr[X \le c] = 1 - K \sum_{r=0}^{n-c-1} \frac{(-1)^r}{r+c+1} \binom{n-c-1}{r} \cdot \left(\frac{\lambda}{\lambda + r + c + 1}\right)^{N\beta}. \quad \text{...(4.5)}$$

It is well known that

$$\frac{1}{r+c+1}\left(\frac{\lambda}{\lambda+r+c+1}\right)^{N\beta}$$

$$= \frac{1}{\Gamma(N\beta)}\int_0^\infty e^{-(r+c+1)t}\,\gamma(N\beta,\lambda t)\,dt .$$

And hence (4.5) can be written as

$$Pr[X \le c] = 1-\frac{K}{\Gamma(N\beta)}\int_0^\infty dt\, e^{-(c+1)t}\cdot$$

$$\gamma(N\beta,\lambda t)\sum_{r=0}^{n-c-1}(-1)^r\binom{n-c-1}{r}e^{-rt}$$

$$= 1-\frac{K}{\Gamma(N\beta)}\int_0^\infty e^{-(c+1)t}(1-e^{-t})^{n-c-1}$$

$$\gamma(N\beta,\lambda t)\,dt$$

$$= 1-\frac{K}{2\pi i}\int_{-i\infty}^{i\infty}\frac{ds}{s}\left(\frac{\lambda}{\lambda+s}\right)^{N\beta}\cdot$$

$$B(c+1-s,\, n-c) . \qquad \ldots(4.6)$$

If $\lambda=1$ and $N\beta$ is an integer, (4.6) reduces to (3.4) which implies that binomial-gamma distribution is equivalent to binomial-uniform distribution.

Thus the following equivalence theorem:

Theorem (4.1) The binomial-gamma distribution with $\lambda=1$ and integer valued $N\beta$ is equivalent to the binomial-uniform distribution with N replaced by $N\beta$.

Following the procedure of section 3, (4.6) can be reduced to

$$Pr[X\le c] = \frac{1}{B(c+1,n-c)}\sum_{k=0}^{N\beta-1}\frac{(-\lambda)^k}{k!}\cdot$$

$$B^{(k)}(c+\lambda+1,\, n-c) . \qquad \ldots(4.7)$$

and the asymptotic representation

$$Pr[X\le c] \sim$$

$$\frac{\Gamma\left(N\beta,\, \lambda \ln\left(\frac{n+1}{c+1}\right)\right)}{\Gamma(N\beta)} ,$$

$$(n, c) \to \infty ,$$

$$= \frac{\Gamma\left(N\beta,\, \ln\left(\frac{n+1}{c+1}\right)^{\lambda}\right)}{\Gamma(N\beta)} . \qquad \ldots(4.8)$$

Let $N\beta=1$, then (4.7) reduces to

$$Pr[X\le c] = K.\, B(c+\lambda+1,\, n-c)$$

$$= \frac{\Gamma(n+1)}{\Gamma(n-c)\Gamma(c+1)}\cdot\frac{\Gamma(c+\lambda+1)\Gamma(n-c)}{\Gamma(n+\lambda+1)}$$

$$= \frac{\Gamma(n+1)}{\Gamma(c+1)}\,\frac{\Gamma(c+\lambda+1)}{\Gamma(n+\lambda+1)} . \qquad \ldots(4.9)$$

For large n and c , the asymptotic formula (4.8) yields

$$Pr[X\le c] \sim \left(\frac{c+1}{n+1}\right)^{\lambda} . \qquad \ldots(4.10)$$

Let $N\beta = 2$, then (4.7) yields

$$Pr[X\le c] = K\left[B(c+\lambda+1,\, n-c) - \lambda B^{(1)}(c+\lambda+1,\, n-c)\right]$$

$$= K\, B(c+\lambda+1, n-c)\left[1+\psi(n+\lambda+1)-\psi(c+\lambda+1)\right]$$

$$= \frac{\Gamma(n+1)}{\Gamma(c+1)}\cdot\frac{\Gamma(c+\lambda+1)}{\Gamma(n+\lambda+1)}\cdot$$

$$\left[1+\psi(n+\lambda+1)-\psi(c+\lambda+1)\right] . \qquad \ldots(4.11)$$

$$Pr[X \le c] \sim$$

$$\left(\frac{c+1}{n+1}\right)^{\lambda}\left[1+\lambda \ln\left(\frac{n+1}{c+1}\right)\right], \quad (n,c)\to\infty .$$

...(4.12)

And similarly, for $N\beta = 3$,

$$Pr[X \le c] = \frac{\Gamma(n+1)}{\Gamma(c+1)}\frac{\Gamma(c+\lambda+1)}{\Gamma(n+\lambda+1)}$$

$$\Big[1+\psi(n+\lambda+1)-\psi(c+\lambda+1)$$

$$+\frac{1}{2}\left\{\left(\psi(c+\lambda+1)-\psi(n+\lambda+1)\right)^2\right.$$

$$\left.+\left(\psi'(c+\lambda+1)-\psi'(n+\lambda+1)\right)\right\}\Big]$$

$$\sim \left(\frac{c+1}{n+1}\right)^{\lambda}\left[1+\lambda \ln\left(\frac{+1}{c+1}\right)+\right.$$

$$\left.\frac{1}{2}\lambda^2\left(\ln\left(\frac{n+1}{c+1}\right)\right)^2\right], \quad (n,c)\to\infty.$$

...(4.13)

5. Relation Between Binomial and Pascal Distribution

Consider the Pascal random variable X, given by the probability function,

$$Pr[X = x] = \binom{x-1}{k-1} p^k (1-p)^{x-k},$$
$$x = k,\ k+1,\ \ldots\ldots$$

$= 0$, otherwise,

...(5.1)

where X is the number of experiments to be performed in order to achieve k successful experiments, and p is the probability of achieving success in a single experiment ($q = 1-p$).

The probability distribution function of X is given by

$$Pr[X \le c] = \sum_{x=k}^{c} \binom{x-1}{k-1} p^k (1-p)^{x-k}, \quad c \ge k$$

...(5.2)

The probability distribution of X given by (5.2) is closely related to binomial random variable with parameters (n, q). This relationship is given by the following equivalence theorem:

Theorem (5.1) Let X be Pascal with parameters (k, p) and Y be binomial with parameters (n, q), $(q = 1-p)$, then the probability distribution functions $Pr[X \le c] \equiv Pr[Y \le c']$, provided $n = c$ and $c' = c - k$.

Proof: From Eq. (5.2)

$$Pr[X > c-1] = \sum_{x=c}^{\infty} \binom{x-1}{k-1} p^k q^{x-k}$$

$$= \binom{c-1}{k-1} p^k q^{c-k}\Big[1+\frac{c}{c-k+1}q$$

$$+\frac{c(c+1)}{(c-k+1)(c-k+2)}q^2+\cdots\Big]$$

$$= \binom{c-1}{k-1} p^k q^{c-k}\,{}_2F_1(c, 1, c-k+1; q),$$

...(5.3)

where ${}_2F_1(a, b, c; z)$ is hypergeometric function.

Using the well known identity

$$(1-z)^{a+b-c}\,{}_2F_1(a, b, c; z)$$

$$= {}_2F_1(c-a, c-b, c; z),$$

Eq. (5.3) can be written as

$$Pr[X > c-1]$$

$$= \binom{c-1}{k-1} q^{c-k}\,{}_2F_1(1-k, c-k, c-k+1; q)$$

$$= \binom{c-1}{k-1} q^{c-k} \sum_{m=0}^{k-1} \frac{(1-k)_m\,(c-k)_m}{(c-k+1)_m}\frac{q^m}{m!}$$

$$= \binom{c-1}{k-1} q^{c-k}\Big[1+(c-k)\cdot$$

$$\sum_{m=1}^{k-1}\frac{(k-1)_m}{m!}\frac{(-q)^m}{c-k+m}\Big],$$

or, $$Pr\,[X > c-1] = \binom{c-1}{k-1}\Big[q^{c-k} + (c-k)\sum_{m=1}^{k-1} (-1)^m \binom{k-1}{m} \frac{q^{c-k+m}}{c-k+m} \Big]$$

$$= \binom{c-1}{k-1}\Big[q^{c-k} + (c-k)\sum_{m=1}^{k-1} (-1)^m \binom{k-1}{m} \int_0^q x^{c-k+m-1}\,dx \Big]$$

$$= \binom{c-1}{k-1}\Big[q^{c-k} + (c-k)\int_0^q dx\; x^{c-k-1} \sum_{m=1}^{k-1} (-1)^m \binom{k-1}{m} x^m \Big]$$

$$= \binom{c-1}{k-1}\Big[q^{c-k} + (c-k)\int_0^q dx\; x^{c-k-1} \left\{ (1-x)^{k-1} - 1 \right\} \Big]$$

$$= \binom{c-1}{k-1}(c-k)\int_0^q x^{c-k-1}(1-x)^{k-1}\,dx$$

$$= \frac{1}{B(c-k,k)}\int_0^q x^{c-k-1}(1-x)^{k-1}\,dx,$$

or $$Pr\,[X \le c-1] = \frac{1}{B(k,c-k)}\int_q^1 x^{c-k-1}(1-x)^{k-1}\,dx. \qquad \ldots(5.4)$$

The probability distribution function of binomial random variable Y with parameter (n, q) is given by

$$Pr\,[Y \le c'] = \frac{1}{B(n-c', c'+1)} \int_q^1 y^{c'}(1-y)^{n-c'-1}\,dy. \qquad \ldots(5.5)$$

From (5.4) and (5.5), it is clear that $Pr\,[X \le c] = Pr\,[Y \le c']$, provided $n = c$ and $c' = c-k$. Thus the compound Pascal distribution can be obtained by the results on compound binomial distribution obtained in sections 1, 2, 3, and 4.

Application to Reliability

6. Reliability of Series-Parallel System

Consider a series parallel arrangement with independent and identical subsystems in parallel. Each subsystem contains m relays in series. The system is successful if at least k out of n parallel subsystems operate. The probability of success of jth relay in the series is p_j. In the model p_j's $(j = 1, 2, \ldots\ldots, m)$ are assumed to be independent random variables with known probability distribution. The problem considered is to determine system reliability.

Let $p = p_1 \cdot p_2 \cdot \ldots\ldots p_m$, p_j's are random variables defined in the interval $(a_j, b_j) \subseteq (0, 1)$, and X be the number of successful subsystems. Then the system reliability is

$$R_s = Pr\,[X \ge k] \qquad \ldots(6.1)$$

$$= Pr\,[X > k-1]$$

$$= \sum_{r=1}^{n-k+1} M_p(k+r) \prod_{j=1,\neq r}^{n+1-k} \left(\frac{j+k-1}{j-r}\right) \text{; from} \qquad \ldots(6.2)$$

(1.6), where $M_p(\alpha)$ is the Mellin transform of p.

Hence
$$R_s = \sum_{r=1}^{n-k+1} \prod_{i=1}^{m} M_{p_i}(k+r) \prod_{j=1,\neq r}^{n-k+1} \left(\frac{j+k-1}{j-r}\right), \qquad \ldots(6.3)$$

since p_j's are independent.

Thus the determination of the system reliability involves the knowledge of the Mellin transform of individuals p_j's.

7. Approximation of Reliability of Series-Parallel System

Beta Distribution

Let p_j, $j = 1, 2, \ldots\ldots, m$, be independent and have beta distribution given by the density function

$$f_{p_j}(t) = \frac{t^{a_j-1}(1-t)^{b_j-1}}{B(a_j, b_j)}, \quad 0 \le t \le 1,$$

$\{a_j, b_j\}$ are large.

The system reliability is given by

$$R_s = P_r[X \ge k]$$

$$= 1 - P_r[X \le k-1]$$

$$\sim K \int_0^{\omega} x^{k-1}(1-x)^{n-k}\,dx, \text{ from (2.5)}$$

$$= K\, B_\omega(k, n-k+1), \quad \ldots(7.1)$$

where

$$K = [B(k, n-k+1)]^{-1}$$

$$\omega = \prod_{j=1}^{m} \frac{a_j}{a_j+b_j}$$

and $B_\omega(p, q)$ is incomplete beta function.

Uniform Distribution

Let p_j, $j = 1, 2, \ldots\ldots, m$, be independent and have uniform distribution in the interval $(0, 1)$.

From (3.5), the system reliability is obtained as follows:

$$R_s = P_r[X \ge k]$$

$$= 1 - P_r[X \le k-1]$$

$$= 1 - K \sum_{j=0}^{m-1} \frac{(-1)^j}{j!} B^{(j)}(k+1, n-k+1)$$

$$= 1 - \frac{1}{B(k, n-k+1)} \cdot \sum_{j=0}^{m-1} \frac{(-1)^j}{j!}$$

$$B^{(j)}(k+1, n-k+1), \quad \ldots(7.2)$$

where $B^{(j)}(x, c) = \frac{d^j}{dx^j} B(x, c)$

If n and m are large, then using the asymptotic result (3.3-6), the system reliability is given by

$$R_s \sim 1 - \frac{\Gamma\left(m, \ln\left(\frac{n+1}{k}\right)\right)}{\Gamma(m)}$$

$$= \frac{\gamma\left(m, \ln\left(\frac{n+1}{k}\right)\right)}{\Gamma(m)}. \quad \ldots(7.3)$$

Gamma Distribution

Let $p_j = \exp(-\omega_j)$, $j = 1, 2, \ldots, m$ and ω_j's be independent and identical random variables having gamma distribution given by the p.d.f.

$$f_{\Omega_j}(\omega) = \frac{\lambda^\beta \omega^{\beta-1} e^{-\lambda\omega}}{\Gamma(\beta)}, \quad \omega > 0$$

From (4.7) the system reliability is given by

$$R_s = P_r[X \ge k]$$

$$= 1 - \frac{1}{B(k, n-k+1)} \sum_{j=0}^{m\beta-1} \frac{(-\lambda)^j}{j!}$$

$$B^{(j)}(k+\lambda, n-k+1), \quad \ldots(7.4)$$

($m\beta$ is positive integer).

If n and m are large, then by asymptotic result (4.8)

$$R_s \sim \frac{\gamma\left(m\beta, \lambda \ln\left(\frac{n+1}{k}\right)\right)}{\Gamma(m\beta)}. \quad \ldots(7.5)$$

8. Application to Manufacturing Process

Consider a manufacturing process involving a series of operations. Let there be m operations in series and a product passes through each operation only once. At each operation there is a "yield y_i" defined as the probability that a product passes through operation i successfully. The y_i is thus the characteristic of the operation i and can be measured by experiment, namely, the ratio of the number of good units at the output of the operation i to the number of units at the input. The defective units at each operation are rejected and there is no rework. The demand for finished goods is known and the operation yields (y_i's) are random variables with known probability distribution. The problem considered is to determine the quantity to start with at the beginning of the first operation so as to meet the demand for the finished goods with specified service level (defined as the probability that the quantity

of finished goods resulting from m operations exceeds the demand).

The effective yield of the entire process involving m operation is given by

$$Y = y_1 \cdot y_2 \cdot \dots \cdot y_m \qquad \dots(8.1)$$

Thus Y is the probability that a product passes through m operations successfully. Let k be the number of finished goods desired, and n be the quantity to start at the first operation. Clearly the number of finished goods X out of n is binomial random variable with parameters (n, y). Since y is random variable, the probability distribution of X is given by the compound binomial distribution studied in section 1.

The service level condition is given by $Pr[X \geq k] = \theta$, (a given constant).

or
$$\sum_{r=1}^{n-k+1} M_Y(k+r) \prod_{j=1, \neq r}^{n+k+1} \left(\frac{j+k-1}{j-r}\right) = \theta. \qquad \dots(8.2)$$

Thus the problem reduces to determine n, from the identity (8.2).

Acknowledgment

The author is indebted to Professor J. H. K. Kao for his advice and inspiration to write this paper.

References

1. Prasad, R. D. "Probability Distribution of Algebraic Functions of Independent Random Variables," SIAM Journal on Applied Mathematics, 18 #3 (1970), pp. 614-626.

2. Springer, M. D., and Thompson, W. E. "The Distribution of Product of Independent Random Variables," SIAM Journal on Applied Mathematics, 18 #3 (1965), pp. 511-526.

3. Dubey, S. D. "Compound Pascal Distribution," Annals of the Institute of Statistical Mathematics, 18 #3 (1966), pp. 357-365.

4. Abraham, C. T., and Prasad, R. D. "A Stochastic Model for Manufacturing Cost Estimating," IBM Journal of Research and Development, 13 (1966), pp. 511-526.

5. Lipow, Myron, and Lloyd, David K. Reliability: Management, Methods and Mathematics, Prentice-Hall, Englewood Cliffs, New Jersey. 1962.

6. Barlow, R. E., and Proschan, F. Mathematical Theory of Reliability, John Wiley and Sons, New York. 1965.

7. Carrier, G. F., Krook, M., and Pearson, C. E. Function of a Complex Variable, McGraw-Hill, New York. 1966.

8. Handbook of Mathematical Functions, Applied Mathematics Series 55, National Bureau of Standards, Washington, D. C. 1964.

9. Bateman, H. Tables of Integral Transforms (vol. 1), McGraw-Hill, New York. 1954.

THE ROLE OF TESTING IN ACHIEVING AEROSPACE SYSTEMS EFFECTIVENESS

Anthony M. Smith
General Electric Company
Philadelphia, Pennsylvania

and

Thomas D. Matteson
United Air Lines, Inc.
San Francisco, California

SUMMARY

The relationship between testing and the achievement of aerospace systems effectiveness was the subject of a recent investigation conducted by the AIAA Systems Effectiveness and Safety Technical Committee. The Committee studied various aspects of test program requirements, philosophy and experience for spacecraft, launch vehicles, DOD aircraft and commercial aircraft. A survey questionnaire, completed by both industry and government personnel, served as the primary source of data. This paper summarizes those data dealing mainly with acceptance testing. In this regard, the majority of the survey respondents indicated that acceptance tests play a less-than-dominant role in the achievement of systems effectiveness. A full report of the study findings are to be published by the AIAA.

INTRODUCTION

This paper is a summary report based on the findings of a study project conducted by the AIAA[1] Systems Effectiveness and Safety Technical Committee.

This Committee, whose goals are in part directed toward the development and communication of pertinent technical information in the System Effectiveness disciplines, elected in 1968 to undertake a project that would address these goals and thereby realize some substantative contribution to the professional aerospace community. The selection of an appropriate subject area for the project was in itself a matter of considerable thought and debate. In fact, some 32 candidates topics were reviewed initially and narrowed to four by a committee sub-group; finally by full Committee action the topic of "testing" was selected. The project was later given the official designation that is carried as the title for this paper.

The investigative portion of this project was conducted via a survey type questionnaire in the 1969 to 1971 time period, and the compilation and editing of results for a report was done in 1972. While the basic input data is thus 2 to 3 years old at this point in time, it is the consensus of the Committee that this information maintains its relevance today (and probably will for some time to come) simply because the state of test technology and methodology is a slowly evolving discipline that has, to a major extent, seen no dramatic changes or shifts in emphasis in recent years.

By way of clarification, the term "System Effectiveness" as used herein refers to the disciplines of Reliability, Maintainability and Safety - and hence, the project attempted to relate the role of testing to these discipline areas. Even with this limitation, it was clear to all that the project was attempting to cover a lot of territory. Obviously, many stones were left unturned - but by the same token, the committee feels that some of the more pertinent issues were addressed. At a minimum, the committee hopes that the project results will motivate the continued development and application of effective test technology in the aerospace and related industries.

The Project scope included the Development, Qualification, Acceptance and Flight test phases for spacecraft, launch vehicles, DOD aircraft and commercial aircraft. In keeping with the session theme developed for this Symposium, however, the paper will primarily emphasize acceptance test in these four product areas - recognizing, of course, in the broader sense that test program optimization must always consider the interrelationships that exist between these aforementioned basic test phases.

A full report on this project will be published by the AIAA, and copies can be obtained through AIAA headquarters in New York City.

PROJECT OBJECTIVES AND APPROACH

Simply stated, this Project was undertaken by the Committee with two distinct objectives in mind:

- To ascertain the "state-of-the-art" of aerospace equipment testing relative to the various philosophies, approaches and practices employed throughout Government and industry - with specific attention to testing as it related to reliability, maintainability and safety.
- To accumulate information that would provide at least some partial description of the technical problems or gaps that existed in the area of test technology.

To acquire such information, the Committee developed a Program Plan which provided for the implementation of an industry-wide survey on test techniques and technology and the collection of pertinent literature references to augment the survey data. The Program Plan further stiuplated that the survey and reference material would address four basic product areas (spacecraft, launch vehicles and both DOD and commercial aircraft), four test phases (development,

(1) American Institute of Aeronautics and Astronautics

qualification, acceptance and flight or in-service tests) and test levels ranging from piece part to total system. Finally, the Program Plan provided for the compilation of the data and its publication via summary papers such as this and a final AIAA sponsored report.

The Program Plan was followed by a detailed Guidelines Document which specified the Questionnaire to be used in the survey as well as the definition of several ground rules for the conduct of the Project. The more important of these ground rules were as follows:

1. To seek information relevant primarily to systems effectiveness parameters - i.e., to avoid an in-depth study of items such as facilities, test equipment, instrumentation, data reduction and processing, etc.

2. To confine the scope of interest to "flyable" hardware, thus eliminating all considerations of GSE testing.

3. To obtain information from a broad cross-section of the aerospace community, thus avoiding any bias toward a single program, organization or unique line of thinking.

4. To avoid the use of any classified or proprietary information.

5. To avoid any identification of specific persons or organizations with specific information (unless, as a published reference, it was so identified).

6. To make technical observations or correlations on the data, but to avoid value judgements per se on the "goodness" of any information, relying on the reader to apply such judgements as he may deem appropriate.

The Questionnaire, the main source of the project data, was composed of 30 main questions. Many questions contained sub-sets of related questions and, all told, there were about 45 questions. The majority of the questions requested an answer in a fixed, multiple-choice format to facilitate the compilation of statistics, but almost every question also requested a narrative explanation to provide assistance in the statistical analysis. About half of the Questionnaires were completed via a personal interview conducted by one of the Committee members; the others were completed by mail with some phone conversation where such was possible to clarify certain responses.

The Committee did attempt to find and record published literature which provided data of interest to many of the survey questions. The respondents, themselves, also provided references as a part of their input. However, an intensive literature search was not conducted nor was any attempt made to make an in-depth correlation between the Questionnaire data and published literature. This may be considered a weakness in this Project, but the practicality of time availability necessitated such a choice. A list of references obtained in the course of this Project is not included here, but will be found in the AIAA final report.

The following section presents some selected survey questions and their results. As noted earlier, this paper attempts to draw upon those questions which relate primarily to the subject of acceptance testing.

PROJECT FINDINGS

Some Preliminary Comments

The data for the findings discussed in this section are based on 56 responses to the survey questionnaire. With but minor exceptions, the responses come from different companies, Government agencies or completely separate Divisions within these organizations - thus achieving the desired effect of a wide cross-section of inputs. The responses are divided among the four product areas of interest as follows:

Spacecraft	-	12
Launch Vehicles	-	15
DOD Aircraft	-	21
Commercial Aircraft	-	8

Somewhat understandably, the commercial aircraft group represents the smallest data sample in the survey. For purposes of analysis, this group has also been divided into equipment manufacturers and airline operators due to the widely differing vantage points from which the respondents addressed the questions. In this paper, the 8 respondents are all equipment manufacturers; the complete final report includes some discussion on 4 additional responses from the airline operators.

At this point, a word of explanation and precaution is in order regarding the reader's interpretation and use of the results. Most importantly, one should remember that data from each respondent within a product area was based on his particular experience and vantage point, and thus responses to any given question can vary widely even though the same type of product, the same type of specifications and the same ultimate customer were involved. In other words, the data reflect what people "think" to be the answer - even though in a more detailed analysis or discussion they might be convinced that the "facts" do not support their perception. Of course, in going across the four product areas, differences of opinion often become even more pronounced. Interestingly enough, this apparent problem is actually one of the strong points of the study simply because people do indeed act (and managers to indeed make decisions) on what they "think", not necessarily on what are always the "facts". In this same regard, the survey did not require corroborative evidence to support each respondent's answer, nor did it request the respondent to obtain higher management approval before submittal (although it is known that the latter occurred in several, if not most, responses).

Also, for reasons not explained, every respondent did not always answer every question - so often, the data sample for a given question may be somewhat less than 56. But the answers provided to the Committee were used as submitted even if they tended to reflect some inconsistency with information provided elsewhere in the response. (Incidentally, this in itself provided some early clues to substantiate the observation of many Committee members that general inconsistencies often appear in conversations about test philosophies and techniques).

In summary, the findings should be viewed as trends within a given product area or as relative points of comparison between product areas. To rely on their value as absolute indicators could be misleading.

Sample Survey Results

Format-wise, selected questions from the survey are presented below as they appeared in the Questionnaire, results are tabulated for each product area to make visible those differences or similarities that exist, and some observations are made on the results.

1. Environmental Acceptance Test Stress

Question 6B: Where environmental test is used, what general levels of environmental stress are employed relative to expected nominal flight conditions as represented by 100% in the following matrix (e.g. for a nominal flight temperature of 100°F and an environmental test temperature of 150°F, enter 150%)?

	Development			Qualification			Acceptance		
	Comp	S/S	Sys	Comp	S/S	Sys	Comp	S/S	Sys
Vibration									
Thermal									
Thermal Vacuum									

Results: The results for this question are shown on Figure 1, and are limited here to Acceptance Test and subsystem/system data only. An obvious item of interest is the rather wide range of vibration and thermal stress levels used in all four product areas. This could be the reflection of a general uncertainty on just what the "right" acceptance stress level is. Also note that, except for Launch Vehicles, the upper end of these ranges often exceed the expected nominal flight levels. Thus, while many people express concern about acceptance stresses greater than 100%, the survey indicates that much acceptance testing is indeed done above the 100% stress value. On average, Spacecraft appear to use higher environmental acceptance stresses than any other product line - at least at the subsystem and system levels of assembly.

2. Electronic Part Screening

Question 10B: In general, what percentage of electronic piece parts are you required to screen and/or burn-in for a specific Project application?

0 - 25% ________

26 - 50% ________

51 - 75% ________

76 - 100% ________

Results: The results for this question are shown on Figure 2. The weighted average values, which quickly show the differences between product lines, were calculated on the basis that the upper value in the selected range is what the respondent intended to signify as the % screened. This assumption is most likely optimistic and these average values may thus be somewhat high. Nonetheless, it is clear that Space Programs do more screening than Aircraft Programs and that Spacecraft have the most part screening - and Commercial Aircraft the least. With this type of data before us, the next question would seem to be "Do we need more or less part screening for tomorrow's systems? While the survey did not directly address such a question, an indirect measure for at least DOD Aircraft can be inferred from the data of Question 14, Figure 5, where the second highest cause of flight malfunctions was piece part failures versus an average part screening value of 42%. This comparison could thus suggest that, with a continued sophistication in military avionics, additional part screening is in order.

3. Test Contribution to Systems Effectiveness

Question 11A: In what order of precedence would you rate the following types of test in contribution to the achievement of systems effectiveness (1 = highest, etc.)?

	Reliability	Maintainability	Safety
Development			
Qualification			
Acceptance			
In-Service			

Results: The results for this question are shown on Figure 3, indicating only the relative standing of acceptance test within the four choices possible. The obvious message in these results is that acceptance test is not felt to be a significant contributor to systems effectiveness factors. One could argue that this is not the message at all; that acceptance test is very important, but by forcing people to rank, it just looks that way in the statistics. This could be the case - but consider that the question was structured deliberately to force such a ranking because, in the real-world, such ranking does occur (say, in competing for a fixed amount of funds). In such a competition, it appears that acceptance tests run a second best to the other types of test with Spacecraft, and worse yet in the other product areas. An interesting point for the reader to ponder and compare with his own experience!

4. Eliminating Design Problems

Question 11C: How important do you consider the types of tests to be in eliminating design type problems in your product?

	Very Important	Possibly Important	Net Important
Development	____	____	____
Qualification	____	____	____
Acceptance	____	____	____
In-Service (Flight)	____	____	____

Results: The results for "Very Important" only are shown in Figure 4. Unlike Question 11A, the respondent was not required to rank in this question, and he could in fact arbitrarily indicate all four types of tests to be "Very Important". It can be seen in the data that this did not occur, and that a true assessment of importance was expressed for each type of test. Knowing the results of Question 11A, it is probably not surprising now to see that acceptance tests again are demoted to a rather insignificant position when it comes to their importance in eliminating design problems. The traditional notion, that Development and Qualification Tests are the only "Very Important" tests for design problem solutions, emerges from this data (although Aircraft Programs do show a rather strong liking for flight test also). How does that compare with your experience - especially if your experience includes a fairly large production program?

5. Failure Cause

Question 14: What do you consider to be the relative frequency of failure cause in your product? If percentage estimates are not available, rank in order (1 = highest).

	% In Ground Test	% In-Service (Flight)
Design Faults		
Workmanship Defects		
Test Errors		
Random Part/Mat'l Defects		
Requirement/Specification Error		

Results: The results are shown in Figure 5. Overall, the one-two punch is Workmanship Defects and Design Faults in that order. (In the details of the data not shown here, Workmanship and Design were usually closely ranked, but significantly ahead of the third place contender.) The correlation of these results with other data is interesting. For example, the Number 1 cause of failures in both ground and flight tests is almost unanimously Workmanship, yet in Question 11A we found that acceptance test (the primary test screen for workmanship defects) was not considered a prime contributor to Reliability. Or as another example, design faults are either the Number 1 or 2 cause of flight test failures in all but DOD aircraft, yet in Question 11C we again found that acceptance tests (the only tests run on every article prior to flight) were not highly considered as a screen for design problems. Could it be that acceptance tests, as they are conducted today, are in truth not all that good? Since acceptance tests are the only tests run on every article, should they play a more prominent role in the scheme of things? Are acceptance tests as poorly thought of, in reality, as this data indicates? While this Project has not answered these questions, it is important, we believe, that it has helped to assure that at least they were asked.

6. Most Valuable Test

Question 16: What single test that you regularly perform do you consider the most valuable for enhancing Systems Effectiveness?

Results: This was a strictly narrative type response, so all answers indicating an acceptance type test (e.g. part burn in, field checkout, etc.) were grouped and taken as a % of all answers given in order to get the data on Figure 6. About all that can be said at this point is that the data continues to reflect the results seen in Questions 11A and 11C - i.e., acceptance tests do not come on strong when words like "prime contributor, very important and most valuable" are used. The various forms in which the questions were asked and the consistency in the answers received does appear to confirm the validity of this observation.

CONCLUSIONS

Rather than repeat or summarize the observations made on the selected questions used in this paper, we believe that a broader summarization of the total survey is more appropriate.

The one central theme that emerges from this survey is that opinions and case histories vary widely within and between the product lines investigated. There is rarely any universal agreement on what to do, how to do it and what is important or unimportant. Trends may be observed (as we saw in this paper), but the reasons for such trends or their correlation with other data are, at best, often speculative or spotty (or outright negative). While testing consumes a large portion of a project's funds, test program requirements, philosophies and techniques are derived mainly from a process that is apparently more "art" than "science".

The survey results clearly point to the need for a more scientific approach to the specification, planning and implementation of test programs. They also indicate the necessity for a rigorous industry-wide attack on the broader analysis and interpretation aspects of test results with the intent of defining test methods, techniques and practices that will yield more return for the test dollar spent.

ACKNOWLEDGEMENTS

The authors, who were the AIAA Systems Effectiveness and Safety Technical Committee Chairmen during the period of this Project, wish to acknowledge - and, in fact, emphasize - that this Project was accomplished as a result of the efforts and participation of some 40 Committee members over the '69 to '71 time frame. They also wish to specifically acknowledge and thank the following individuals who personally undertook the job of compilation, integration and editing of the results in the four product areas:

Dr. Jack Carter, Martin-Marietta - Spacecraft

Mr. Tod Braunstein, United Technology Center - Launch Vehicles

Mr. James Sindt, Honeywell, Inc. - DOD Aircraft

Mr. Thomas Matteson - Commercial Aircraft
United Air Lines, Inc.

Finally, to those who took the time, thought and effort to answer the Questionnaire, we offer our sincere appreciation.

TYPE OF STRESS	S/C	L/V	DOD A/C	COMM A/C
VIBRATION	70-150%	30-100%	20-160%	< 100-125%
THERMAL	100-300%	30-100%	20-100%	(2)
THERMAL VACUUM	100%	50% (1)	100%	(2)

(1) - NO SYSTEM LEVEL TESTS INDICATED.
(2) - RESPONSES WERE BLANK - POSSIBLY INDICATING NONE.

Figure 1. Question 6B-Levels of Environmental Stress Used in Acceptance Test (S/S and System Level Only)

% SCREENED	SPACE PROGRAMS		AIRCRAFT PROGRAMS	
	S/C	L/V	DOD A/C	COMM A/C
0-25%	18%	55%	60%	100%
26-50%	9%	18%	20%	0%
51-75%	0%	0%	13%	0%
76-100%	73%	27%	7%	0%
WEIGHTED AVERAGE	82%	55%	42%	25%

Figure 2. Question 10B-Screening and/or Burn-in of Electronic Piece Parts

SYSTEMS EFFECTIVENESS PARAMETER	S/C	L/V	DOD A/C	COMM A/C
RELIABILITY	2	3	3	4
MAINTAINABILITY	2	4	4	4
SAFETY	2	3	4	3

Figure 3. Question 11A-Acceptance Test Contribution to Systems Effectiveness (As Ranked Among Development, Qualification, Acceptance and Flight Tests)

TYPE OF TEST	S/C	L/V	DOD A/C	COMM A/C
DEVELOPMENT TEST	75%	93%	95%	100%
QUALIFICATION TEST	82%	86%	55%	71%
ACCEPTANCE TEST	27%	0%	25%	29%
FLIGHT OR IN-SERVICE TESTS	13%	17%	52%	57%

Figure 4. Question 11C-Importance in Eliminating Design Problems, "Very Important" Replies

GROUND TEST

CAUSE	S/C	L/V	DOD A/C	COMM A/C
WORKMANSHIP DEFECTS	1	1	1	2
DESIGN FAULTS	2	2	2	1
PART/MAT'L DEFECTS	4	3	3	4
TEST ERRORS	3	4	4	3
REQ'T/SPEC. ERROR	5	5	5	5

IN-SERVICE

CAUSE	S/C	L/V	DOD A/C	COMM A/C
WORKMANSHIP DEFECTS	1	1	1	1 (TIE)
DESIGN FAULTS	2	2	4	1 (TIE)
PART/MAT'L DEFECTS	3	3	2	3
TEST ERRORS	4	4	3	5
REQ'T/SPEC. ERROR	5	5	5	4

Figure 5. Question 14-Frequency of Failure Cause in Order of Occurrence (1 = Highest, etc.)

	S/C	L/V	DOD A/C	COMM A/C
% REPLIES FOR ACCEPTANCE TYPE TESTS	8%	27%	14%	0%

Figure 6. Question 16-Most Valuable Test for Enhancing Systems Effectiveness-Acceptance Test Replies

THE ROLE OF TEMPERATURE IN THE ENVIRONMENTAL ACCEPTANCE TESTING OF ELECTRONIC EQUIPMENT*

INDEX SERIAL NUMBER – 1044

R. W. Burrows
Martin Marietta Aerospace
Denver, Colorado

Summary

This paper discusses the role of temperature cycling in the acceptance testing of production electronic assemblies and multilayer printed circuit boards, and in the verification of the basic packaging methods used for electronic components. The paper also summarizes the techniques that should be employed to achieve reliable electronic hardware.

Temperature Cycling of Electronic Black Boxes

Industry Survey

Data from 26 aerospace-oriented companies, or agencies (designated herein by the leters A through Z), are summarized in Table 1. It can be seen that the industry practice ranged from one to 25 temperature cycles, with a gross average of about eight cycles. The most commonly used temperatures were -65°F (-54°C) and 131°F (55°C).

Types of Defects

Examples of the types of defects screened out by temperature cycling are:

- Faulty capacitors, transistors, diodes, integrated circuits, etc.
- Shorts and opens in transformers and coils.
- Faulty solder and weld joints.
- Shorts in cabling.
- Faulty insulation washers, lugs shorted to ground, etc.
- Defects in printed circuit boards.
- Problems due to incorrectly applied conformal coating.
- Drift problems.
- Failures of plastic-encapsulated parts.
- Defective potentiometers, relays, etc.
- Improper staking of tubing coil slugs.

Test Philosophy

In the last few years there has been an increasing use of the philosophy that acceptance tests should be designed for maximum practicable effectiveness. As a result, test conditions are often more stringent than the actual flight environment. Qualification testing should include temperature cycling at levels about 20°F (11°C) hotter and 20°F (11°C) colder than those for acceptance testing, and may be essentially determined by the selected acceptance test level and not by the actual flight environment.

In detecting defects by temperature cycling, the equipment should be operated and closely monitored during the testing. However, it is desirable to turn the equipment off during the cooldown portion of each cycle to prevent self-generated heat from keeping the internal parts warm.

Failure Categories

Averaged estimates from eight companies (Table 2) indicated that most failures during temperature cycling of mature hardware fell into three broad categories:

1) Failures due to marginal design - 5%
2) Failures due to fabrication workmanship - 33%
3) Failures due to faulty parts - 62%

In immature hardware, a greater incidence of design failures can be expected. In programs using very thoroughly screened parts, a lower incidence of parts failures would be expected.

Specific Failure Data

Seven of the 26 companies provided specific failure data (Table 3 and Figure 1). These data indicate that six to 10 temperature cycles are required to detect the majority of the defects and to approach the constant-failure-rate portion of the curve.

Specific failure rate data are shown in Figure 2. These data, normalized to the electronic parts count and shown as Figure 3, provide a baseline from which test failure risks and repair costs can be estimated. This figure also emphasizes the reliability problems inherent in equipment of very high complexity and will, hopefully, influence the reader toward systems of lower complexity.

Effect of Low-Level Vibration

Some of the data from the seven companies came from programs using AGREE testing in accordance with MIL-STD-781B, in which equipment is also exposed to 2-g vibration. Every company that used this approach felt that the temperature cycling precipitated from 90 to 95% of the failures, and that the 2-g vibration played a very minor role. This finding is consistent with an investigation by NASA-MSC that concluded that vibration, to be an effective screening tool, should be conducted at levels equal to, or exceeding, 6 g (rms).

Rate of Temperature Change

Typical rates of change of internal parts in black boxes during temperature cycling are shown as curves 3, 4, and 5 in Figure 4. A higher rate of change provides more powerful screening but this issue is quite controversial. Some companies remove covers to achieve a higher rate of change, while other companies avoid high rates of change as unrealistic. This author favors rates of change as typified by the area between curves 3 and 4. These rates are significantly less severe than the rates used during the qualification and acceptance testing of the individual electronic piece parts (curves 1 and 2), but are more severe than those used in normal practice (curve 5).

Failure Criteria

When multiple temperature cycling is used as an acceptance test, it is standard practice to allow repairs without requiring a repeat of the entire test. Some

*The data presented were selected from two of the 26 studies performed under NASA-MSC Contract NAS9-12359, *Long-Life Assurance Study*. Mr. J. B. Fox was the NASA Technical Monitor and R. W. Burrows was Program Manager.

programs have required no failure-free cycles; some have required the last one or two cycles to be failure-free; and one program (involving very simple hardware) required 20 consecutive failure-free cycles. Figure 3 shows that as the hardware becomes complex (more than several hundred parts), passing a 10-cycle test without a single failure approaches a statistical improbability. Since a typical device contains several thousand parts, it is recommended that one final failure-free cycle be required to provide confidence in any prior repair. However, if the repair is not very easily implementable and inspectable, additional failure-free cycles should be considered, if appropriate to the individual case.

Effect of Hi-Rel Parts

The temperature cycling of black boxes should produce fewer failures when Hi-Rel parts are used, although other variables have apparently masked this effect in the data of Figures 2 and 3. For maximum reliability, both Hi-Rel parts and extensive temperature cycling are recommended. Company H strongly recommended 10 temperature cycles when Hi-Rel parts were used, but 25 cycles when Hi-Rel parts were not used. Company K believes it more cost-effective to use JANTX parts, plus the temperature cycling of assemblies, than to pursue an ultra Hi-Rel part program without the temperature cycling of assemblies. Several companies have achieved reliable hardware using minimal parts screening, but these firms customarily require extensive temperature cycling of assemblies.

Relationship Between Multiple Temperature Cycling and Thermal Vacuum Testing

Since this paper recommends the use of increased temperature cycling at the black-box level, the question arises as to whether multiple temperature cycling, say 10 cycles, should be accomplished in a thermal-vacuum test, or whether the two tests, temperature cycling and thermal-vacuum, should be conducted separately. Our recommendation is that they should be conducted as *separate* tests. Our reasons behind the decision are as follows.

Thermal-vacuum testing, in order to effectively assess outgassing phenomena, must consist of *long* soaks at both low and high temperatures, with emphasis on the long-duration, high-temperature soak. But one cycle of a thermal-vacuum test may take days in a costly thermal-vacuum chamber since heat transfer is accomplished by radiation. Extending the duration to, say 10 cycles, would be both very time-consuming and very expensive. In addition, the temperature ramps are quite slow--too slow for the most efficient detection of incipient failures.

Another factor is that in a systems-level thermal-vacuum test, the temperature levels may be too mild for detecting incipient failures. For example, if a spacecraft has a thermal control system, the prime objective of the thermal-vacuum test would be to demonstrate proper performance of the thermal control system, and the individual black-box temperature excursions may be quite mild.

Consequently, it appears much more desirable to conduct multiple temperature cycling on the black boxes using conventional, ambient-air temperature chambers, and to follow this with a conventional thermal-vacuum test.

Degrading Effects of Temperature Cycling

Temperature cycling with *good* parts and *good* packaging techniques is not degrading, even with several hundred cycles. One company has conducted tests out to 300 temperature cycles without any indication of an *increasing* rate of failure. However, the packaging design must be compatible with the temperature cycling program or the acceptance test yield will be reduced. Some situations in which electronic hardware may be adversely affected by temperature cycling are listed below.

1) Solder joints may crack due to inadequate stress relief. One typical problem is the problem with conformally coated transistor cans on spacers when lead stress relief is not provided. This situation also occurs in relays, transformers, and large modules when the studs or pins are soldered into printed circuit boards without provisions for stress relief of the solder joint.

2) Thick applications or heavy fillets of conformal coating can break or damage parts and solder joints. Bridging of conformal coating under flat-bottomed parts is particularly catastrophic and must be avoided.

3) The use of an encapsulating compound with a high modulus of elasticity and high coefficient of thermal expansion may damage parts and connections.

4) Weak parts, such as glass diodes, must be protected by sleeves before applying conformal coating.

5) Plastic-encapsulated parts are frequently a problem in a temperature cycling environment, because of stresses from thermal expansion incompatibilities.

6) Multilayer printed circuit boards may fail by cracking at the plated-through holes if the hole plating is too thin or is not ductile, or if the holes have not been cleaned prior to plating.

The above situations can all be avoided by using good parts and proper packaging techniques, and by using temperature cycling to verify the packaging configuration. These subjects are discussed in subsequent sections of this paper.

In general, electronic parts are not subject to significant degradation from temperature cycling, but there are always exceptions. A recent problem was encountered with a photodiode in which the internal construction contained fine wires encapsulated in epoxy: failures resulted because the metal and plastic had incompatible thermal expansion characteristics.

Extensive investigations by the NASA-MSFC Solder Committee concluded that any good solder joint can tolerate 200 severe temperature cycles from -67°F (-55°C) to 212°F (100°C) without evidence of the start of cracking.

Investigations by RADC and IBM place the state-of-the art of good multilayer printed circuit boards at between 200 and 1000 temperature cycles.

Remarks on AGREE Testing

Some of the data in this paper were derived from programs using AGREE testing in accordance with MIL-STD-781B. The AGREE cycle combines temperature ramps, temperature soaks, and low-level (2-g) vibration. The consensus of the 26 companies surveyed is that the temperature soaks and the low-level vibration play a minor role, and that the AGREE technique is essentially equivalent to a temperature cycling test, with the screening strength of the test mainly dependent on the temperature range, the temperature rate of change, and the number of cycles.

The traditional AGREE approach, in which the test data are employed to demonstrate a required level of reliability, is an extremely powerful forcing function in achieving reliability. Usually, however, it is most cost-effective on high-volume production programs, and has not been widely used on the small production programs typical of much aerospace business. When these types of programs cannot afford AGREE testing, then multiple temperature cycling is an excellent, lower-cost alternative.

Temperature Cycling of Multilayer Printed Circuit Boards

The life of a multilayer printed circuit board (MLB) depends on the capability of the board to withstand the temperature-induced stresses resulting from both the soldering process and the subsequent temperature cycling experienced during this service life as a result of both ambient temperature changes and the temperature changes induced when the equipment is energized. The prime failure modes--cracking in the barrel and at the corners of the plated-through holes--are strongly influenced by the different thermal expansions of copper and glass-epoxy. The life of an MLB electroplated with brittle copper, and with thin plate in the plated-through holes, is extremely short since failure will occur after a very few temperature cycles. Through good design and process control, the life of an MLB can be extended beyond 200 temperature cycles between -85°F (-65°C) and 230°F (110°C).

The prime factor in achieving a long-life MLB is the ductility of the copper plate in the plated-through hole. Extreme process control is required during the electroplating process to avoid brittle or hard copper and to ensure high ductilities in regions where the elongation can reach 5 to 10%. The hole drilling and cleaning processes are almost as critical, and also require very close process control. The design is another important factor: hole plating should be at least 0.0015 inch (0.004 cm) thick for long-life applications. Multilayer boards are critical since they are costly and essentially unrepairable.

As a final verification of both the design and the process controls, a test coupon from each production board should be subjected to temperature cycling between -85°F (-65°C) and 230°F (110°C). This coupon should contain 80 to 100 plated-through holes in series. During temperature cycling, any increased, out-of-spec electrical resistance would constitute a failure. The number of temperature cycles should be determined from an analysis of the particular program or mission. RADC has recommended 50 temperature cycles for nominal usages, and this value is currently being employed on the Viking Lander program.

Temperature Cycling Verification of the Packaging Technique

It has been previously stated that temperature cycling, as employed in the acceptance testing of electronic black boxes, is not degrading when good parts and packaging techniques are used. Past history on the Saturn and Apollo programs has indicated that serious generic problems are more apt to result from poor packaging techniques than from defective electronic parts.

Serious problems of solder joint cracking have usually resulted from inadequate stress relief at the solder joint. Temperature cycling of an assembled printed circuit board, containing parts without stress-relief provisions, can produce cracked solder joints in relatively few cycles. The mounting of transistors and larger components, such as multipin modules, relays, and transformers, requires particular attention. To avoid serious and very expensive corrective action, the electronic packaging should be controlled by a packaging specification such as NASA's MSFC-STD-136, *Standard Parts Mounting Design Requirement*. This document, originated by the MSFC Solder Committee, is currently gaining acceptance at other NASA centers. It provides detailed guidelines and drawings of preferred parts-mounting configurations designed to eliminate stress on the solder joint -- a fundamental requirement for reliable electronic hardware. It is important to assure that an ample packaging envelope is selected so that there is space available to employ the stress relief provisions of MSFC-STD-136.

Heavy coats of conformal coating, heavy fillets, and bridging or conformal coating under flat-bottomed parts may also break both the parts and the solder joints in a few temperature cycles. Minimum thicknesses (a few thousands of an inch) should be applied to avoid bridging and heavy fillets.

Parts and connections within potted modules generally experience very high internal pressures, and can be damaged by temperature cycling unless the encapsulating material is carefully selected to avoid this problem.

To verify the adequacy of the basic packaging technique, we recommend testing to MSFC-STD-136, which also requires that the packaging technique be verified by 200 temperature cycles from -67°F (-55°C) to 212°F (100°C). This test basically constitutes an acceptance test of the packaging design and must be conducted in the *early* phases of prototype development, long before qualification and acceptance testing of the developed hardware.

Conclusions

To achieve reliable electronic hardware, temperature cycle testing should be emphasized in three key areas:

1) In verifying the adequacy of the basic packaging technique. This testing should be conducted on early prototypes -- well in advance of qualification and acceptance testing -- and should be considered as the acceptance test of the basic packaging design.

2) In the acceptance testing of printed circuit boards, particularly multilayer printed circuit boards, to ensure that the plated-through holes will not crack and cause an electrical failure.

3) In the acceptance testing of production black boxes to detect incipient failures from marginal design, defective parts, and faulty workmanship.

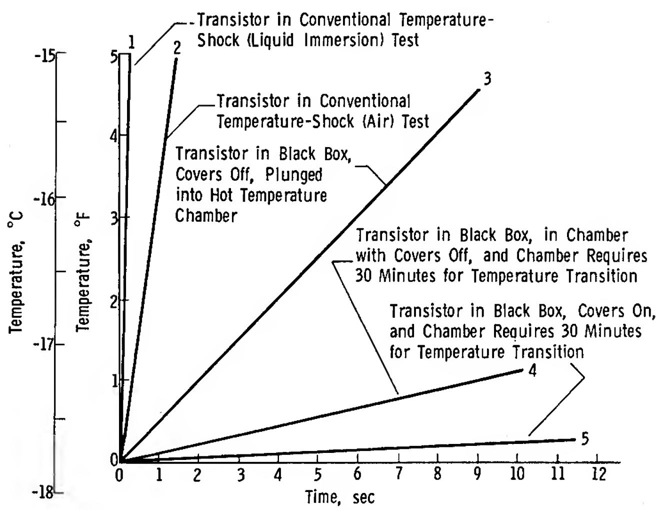


Fig. 4 Comparison of Temperature Change Rates in Various Types of Temperature Testing

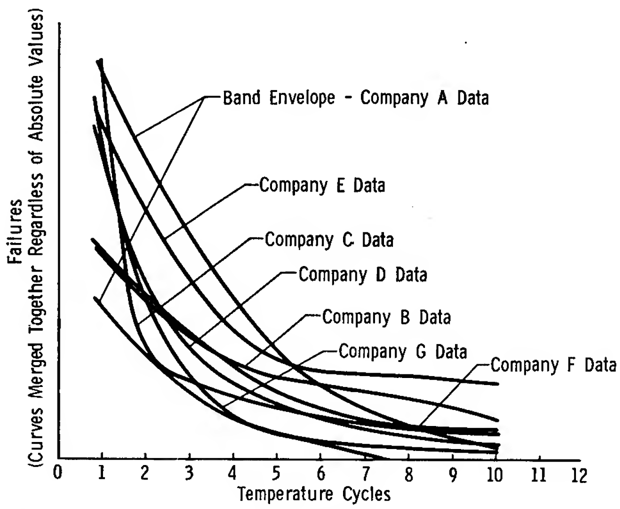


Fig. 1 Summary of Industry Survey Data, Showing That Six to Ten Cycles Are Required for Elimination of Incipient Defects

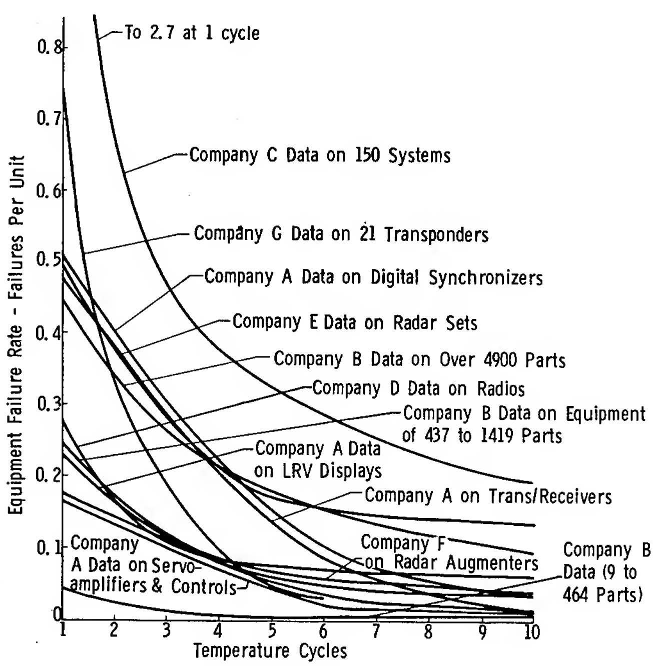


Fig. 2 Summary of Industry Failure Rate Data

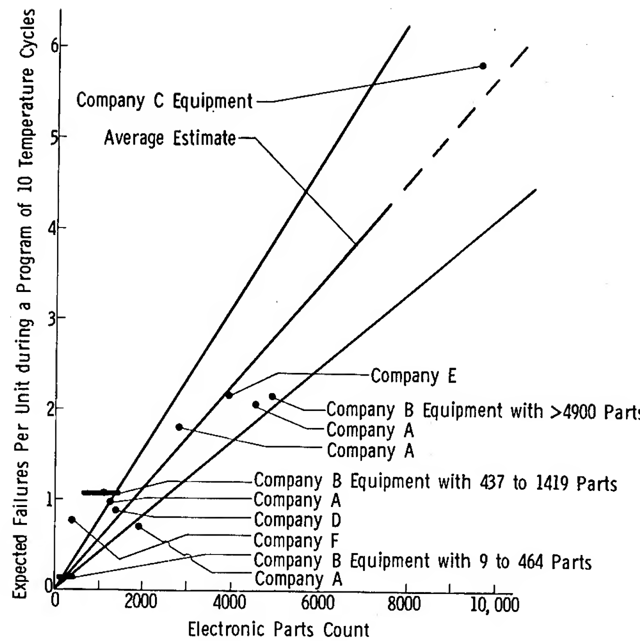


Fig. 3 Expected Failures During Ten Temperature Cycles as a Function of Equipment Complexity

Table 1 Summary of Recommended Temperature Cycles from Industry Survey

Company	No. of Cycles Recommended	Temperature Employed, °F	Temperature Employed, °C	Temp Range, °F	Temp Range, °C
A	6 to 10	-65 to 131	-54 to 55	196	91
B	8 to 10	-20 to 160	-29 to 71	180	82
C	3 to 12	-65 to 131	-54 to 55	196	91
D	9 to 25	-65 to 160	-54 to 71	225	107
E	20	5 to 131	-15 to 55	126	52
F	22	-65 to 160	-54 to 71	225	107
G	6 to 8	Variable	Variable	--	--
H	10 to 25	-65 to 131	-54 to 55	196	91
I	8	Variable	Variable	--	--
J	6 to 10	Variable	Variable	160	71
K	Variable	Variable	Variable	--	--
L	12	-13 to 131	-25 to 55	144	62
M	16	32 to 131	0 to 55	99	37
N	4 to 6	Variable	Variable	--	--
O	6	Variable	Variable	--	--
P	5	-20 to 120	-29 to 49	140	60
Q	5	32 to 160	0 to 71	128	53
R	3	Variable	Variable	--	--
S	3 or 4	Variable	Variable	--	--
T	3 to 5	-65 to 160	-54 to 71	225	107
U	2 to 10	-67 to 131	-55 to 55	198	92
V	2 or more	Variable	Variable	--	--
W	1	Variable	Variable	--	--
X	1	Variable	Variable	--	--
Y	5	-65 to 131	-54 to 55	196	91
Z	1	-65 to 165	-54 to 74	230	110

Table 2 Distribution of Temperature Cycling Failures Between Marginal Design, Poor Workmanship, and Defective Parts

Company	Maturity of Hardware	Percentage of Failures by Categories		
		Design	Fabrication Workmanship	Parts
A	Immature	33%	33%	34%
	Mature	Approaches 0	10%	90%
B	Immature	10%	50%	40%
	Mature			
C	Mature	Approaches 0	50%	50%
D	Immature	33%	33%	34%
	Mature	25%	25%	50%
E	Mature	5%	40%	55%
F	Mature	Approaches 0	10%	90%
J	Mature	5%	35%	60%
L	Mature	Approaches 0	40%	60%
Averages for Mature Equipment		5%	33%	62%

Table 3 Companies Supplying Failure Rate vs Temperature Cycle Data

Company	Type and Size of Data Sample	Were Hi-Rel Parts Used?	Was Vibration a Part of the Cycle?	Temperatures Employed
A	80 Radar Systems	Yes	No	-65°F to 131°F (-54°C to 55°C)
B	80 Command Control Systems	Yes	No	Variable - Most temperature differentials were 160°F (71°C)
C	150 Electronic Systems	Yes	Yes (2 g)	-65°F to 131°F (-54°C to 55°C)
D	360 Radios	No	Yes (2 g)	-65°F to 131°F (-54°C to 55°C)
E	10 Radar Systems	No	No	5°F to 131°F (-15°C to 55°C)
F	270 Radar Augmenters	No	Yes (2 g)	-65°F to 160°F (-54°C to 71°C)
G	21 Transponders	Yes	Yes	Unknown

COST EFFECTIVE HIGH RELIABILITY IN RADIO REMOTE CONTROL SYSTEMS

Michael L. Gilbertson
Technical Director
Kraft Systems, Inc.
450 West California Avenue
Vista, California 92083

The design and manufacture of radio remote control systems for model aircraft requires special methods to achieve the high reliability demanded by the application. Actually a consumer product, these radio control systems must provide extremely low failure rates at very low cost. In order to achieve the utmost performance at competitive prices, special techniques of procurement and testing are necessary. Thoroughly screened high-reliability components are too expensive, and the small size and weight of the airborne system, some weighing less than 11 ounces, precludes special protection from the 50g vibration levels, oil and fuel saturation, and impact at speeds over 100 M.P.H.

Obviously, this type of reliability must be designed into the product. Testing is necessary,yet most radio control equipment is never test flown prior to the customer initially using the system. The reliability is achieved not by testing each component in excess of intended use. The fundamental concept which has been proven in countless designs and hundreds of thousands of hours of use is to use components and assembly methods which are inherently suited to higher reliability. Yet in most cases, this does not mean Mil qualified components or assembly methods. No customer could afford the equipment if it did. How does a manufacturer buy high reliability parts at a price competitive with standard consumer devices, particularly when there are special needs for certain parameters which only the "Mil-spec" devices seem to guarantee? It is actually quite simple. It does require effort and cooperation, but the results can be more than worth the time spent early in the design cycle.

0.001% FAILURE RATE - 11¢ EACH

The most used and most complex component in a radio control system is the semiconductor. Transistors, diodes, and integrated circuits are the heart of most any electronic equipment produced today. There are so many manufacturers and an endless selection of devices to choose from, it is nearly impossible to select the one device which has all the right specifications to suit the design, particularly when there is a requirement for very high reliability. The most obvious choice is a thoroughly screened Mil qualified device or exhaustive testing, right?

It shouldn't be. Exhaustive screening and special devices should actually be a last resort. Where then, do you find inherently low failure rates? Look at the semiconductor manufacturer. What is his most reliable device? It is the high volume device--the consumer plastic type, in particular. Millions of devices are produced each year in hundreds of device types and yet these devices are generally thought to be the <u>least</u> reliable. It would logically seem not to be true, yet it is.

As well as being reliable, these devices are inherently quite tightly specified. A look at the spec sheet would seem to show otherwise, and yet there is a spec sheet most engineers don't see which really tells the true story. The report the manufacturer uses for process control is one of particular interest. Although it is called by a variety of terms, one most commonly used is "Parametric Distribution Data." Simply stated, a "lot" of devices is characterized with each parameter's actual value expressed as a percentage of that lot. In many cases 90-95% of the devices in the lot will fall into 20% of the allowable limits for that parameter. The point should be obvious. If parameter distribution data is used to your advantage, very closely specified devices can be had at very little cost. The manufacturer can select that 5-10% or even 40% in some parameters for pennies over catalog price if it is strictly a parametric specification. Of course, the device must meet all other requirements, but there are many standard devices which are basically excellent choices in most applications.

Operation in severe temperature extremes can often be much more troublesome. Special devices may be necessary, but here too a reasoned approach must be taken. Examination of the <u>actual</u> environmental characteristics should be done before over-specifying for conditions which may not really exist.

One particular transistor I am familiar with has an established reliability level of better than .001% average failure rate with some 4 million device-hours on record in our facility alone. The cost--less than 11¢ each.

There are other types of devices which are used to which these same basic principles can be applied.

A careful look at what the vendor has the most experience and confidence in is often a better clue to the inherent reliability than the spec sheet. The road to this information is easy to follow since it begins with the vendor's representative--the fellow who calls on Purchasing trying to "sell his wares" to someone who may or may not know exactly what it is he's really buying. He may merely pass along a spec sheet to Engineering or to the Technical Library which is never seen. If the engineers do see the "rep," too often the engineer feels the representative doesn't really know what it's all about. Of course, he doesn't. He doesn't have all the answers, but he is the door <u>to</u> the answers. The advantages of cultivating and working with the vendor's salesman cannot be understated. His access to the factory is almost limitless, and in most vendor plants, he can get copies of reliability data, parametric data, test methods, and other information you need to make an intelligent decision rather than using only a not altogether realistic spec sheet. He is also the first contact when a special design or selected part is necessary. <u>He</u> may not have all the answers, but the

vendor does. The salesman should be, and generally is, more than willing to get any information you need should a special or selected part be necessary.

It is often advantageous to have even standard parts "customized." Kraft Systems makes use of specially marked standard parts where replacement may be made in the field by other than factory authorized service personnel. Several parts have been chosen on the basis of the vendor's parametric data, and another vendor's identical device number may not be suitable in all cases. The special marking prevents undesirable substitutions, and assures factory authorized components in critical locations. The cost of special marking is nominal in even small quantities. In lots of a few thousand, it is generally available at no additional cost.

WHEN NOTHING ELSE WILL DO

There are times when only a specially designed part will do--a special environmental or installation problem which can be solved only by a completely different design. But before trying to become an expert on semiconductor design or spending hours over a drafting table, contact the vendor reps and explain the need as closely as possible. I have found that many vendors have quite a catalog of special devices developed internally or for other customers which aren't always advertised. Utilizing a prior design, even with slight modification can literally save thousands of dollars in time and money spent.

For example, Kraft Systems was looking for a more rugged nickel-cadmium battery for use in the airborne portion of the system. We knew the shortcomings of what we were using, and thought that by having a rugged battery designed, field failures could be minimized. During the initial design phase, the vendor representative contacted the factory about a cell which could withstand the high vibration levels. One of the vendor's engineers remembered a special failure resistent cell designed for a large chain-saw manufacturer. Several samples were constructed, and the field tests were very successful. After more than one year of use, these cells have achieved a virtual zero failure rate in the field compared with a 2-3% failure rate for the cells previously used. The cost of these cells was actually less since they were easier for the vendor to assemble. The start-up cost was zero, since all the design had been done previously.

This doesn't mean I advocate making do in all cases. Kraft Systems has several custom designed parts in normal use. However, the point is to avoid jumping into special high reliability, custom components before all avenues have been investigated.

PUTTING IT ALL TOGETHER

Detailed explanations of assembly methods are somewhat superfluous since each design has its own special assembly criteria. The intended application often dictates the functional layout and basic physical placement of various components. In Kraft Systems' case, the primary requirement is shock and vibration resistance in installations affording very little protection against either. Since size and weight are major considerations, the equipment must be very compact. Impact resistant nylon cases are used throughout and heavy epoxy-glass circuit boards are mounted parallel to the axis of greatest expected shock. Due to the necessary component density, most small parts are mounted vertically. This can cause vibration induced failure if the component leads are not adequately mechanically secured prior to soldering. Significantly higher failure rates were experienced when component leads were simply inserted into board holes over the failure rate when leads were bent flush against the board.

Components themselves must be proven mechanically sound, as well. Here again, especially with semiconductors, those the vendor manufactures in greater quantity have shown to have the best mechanical integrity. The use of potting compounds is generally avoided for a number of reasons. Although silicone compounds can add some shock protection, experience in the R/C field has shown its use to have more drawbacks than advantages.

For example, these compounds exhibit a "domino destruction" tendency where pressure on one part may damage an adjacent component. This occurs mainly with closely spaced vertical components such as resistors.

Also, the use of such compounds often hides damage which may reveal itself only under certain types of vibration. If damage results due to impact and it is repairable, the potting compound prevents adequate inspection, and complicates the removal of defective components.

There is a need for protective coatings, however. Since exposure to corrosive fuels and oils is quite normal, plastic based conformal coatings are widely used to protect the metallic portions of the circuitry from contamination by these substances as well as the occasional salt water "bath" in areas near the ocean.

TESTING

While incoming component inspection is rarely performed in the R/C industry, final testing is thoroughly exhaustive. The approach which dictates this course of action is simple. If it's going to fail, it usually fails in the field. A variation of Murphy's Law, no doubt.

The final test phase begins as each portion of the unit becomes initially operational. The assembly is tested and test results noted on paperwork which remains with the unit. Then the assemblies are placed on a full functional burn-in for a minimum of 24 hours. Following burn-in, any parameter changes are noted. Any unit showing gross change is held for further evaluation or further burn-in.

The procedure was originally adopted to pre-age certain components, especially quartz crystal oscillators and tuning circuits, as well as weeding out weak semiconductors. Since implementation, however, a number of the isolated types of failures have been prevented from reaching the field.

Field failures have declined markedly since initiation of this procedure, as well as providing a double check on each unit's performance. At the post-burn-in test, each unit is functionally monitored while being subjected to temperature/humidity extremes. Any defective components are replaced and the entire

assembly is returned to burn-in.

Once the system has been final tested and is ready for shipment, it is placed in a holding area for a special pre-shipment final inspection. In this area, two things may take place. After remaining 24-48 hours, a final inspection is made by other than final test technicians, to check overall system performance. Approximately one unit in 50 will be returned for further testing, due simply to human error by final test technicians on perhaps one of over 100 inspections required prior to reaching this point.

In some cases, Engineering will pull a unit at random prior to final inspection to evaluate overall the entire system. Tests beyond those in final test are performed mainly to verify the design variables and evaluate production processes. Since Engineering is responsible also for much of the process control as well, evaluations of final test systems are also made "to keep everybody honest," so to speak. Often trends unnoticed in final testing can help prevent a small crisis later when tolerances have slipped too far.

The results of the basic program above have been very impressive. Implemented almost two years ago in a complete sense, field failures have fallen drastically. Previously the return rate on new equipment averaged 5-8% depending on how new the design was. The latest return data indicates a field failure rate of less than 1% for any cause.

All of the design and testing methods described above, while really a common sense approach to the problems of producing a good product, aren't always the most obvious. Most were born of the need to achieve low failure rates at low cost, yet these same methods are applicable to almost any situation.

A side benefit of this program has been a strong team approach on the part of everyone involved--Engineering, Purchasing, Final Test, Production, and our Authorized Service Stations throughout the world. Each individual is encouraged to participate and make suggestions, even though it may be outside their field of responsibility. The line of communication between Engineering and Purchasing is very strong and both benefit, since Purchasing is much more familiar with the components and specifications needed, and can make worthwhile suggestions to Engineering for alternate components when necessary.

The Authorized Service Stations have provided invaluable feedback on field problems and failures. Detailed reports on repairs are sent to the factory for analysis, and have helped identify problems peculiar to certain geographic areas.

The best measure of the program's success has been the number of specific items which have disappeared from failure statistics. Several other minor problem areas are now being investigated, which will increase overall reliability even further. A program like this must be continually in operation to assure the reliability achieved thus far.

One specific problem area under consideration is the feedback potentiometer in the servo. Wear and noise have been nagging problems for several years, although not major problems. The apparent solution is not really in the spirit of the program, however, since it is a completely new plastic potentiometer design rather than an adaptation of an existing design. But as I said earlier, "there are times when only a specially designed..."

AEGIS AN/SPY-1 RADAR SYSTEM - DESIGN FOR AVAILABILITY

INDEX SERIAL NUMBER – 1046

William J. O'Leary
RCA, Missile and Surface Radar Division
Moorestown, N. J.

Summary

The AEGIS Weapons System is a Navy defensive missile system designed to shield the fleet against airborne threats of the late 1970's and 1980's. At the heart of this system is the automatically controlled AN/SPY-1 radar, which provides target search and track and missile guidance functions. Considering its almost one million parts the AN/SPY-1 radar presents the greatest challenge to the achievement of AEGIS availability design objectives. Through specific examples, this paper shows how RCA approached a variety of reliability and maintainability design problems relating to the AN/SPY-1 radar and the overall success achieved. Availability design solutions are related to system support characteristics, such as manning and sparing. This relationship is shown in a requirements structure which relates system effectiveness elements. These elements are categorized into four major groupings with specific design objectives for each. It is shown that the AN/SPY-1 design solution achieves not only the high levels of availability desired but also has excellent support characteristics.

Introduction

AEGIS is a Navy defensive missile system designed to shield the fleet against airborne threats of the late 1970's and 1980's. Figure 1 illustrates AEGIS ships deployed as escorts in an engagement situation. At the heart of the AEGIS Weapons System is its automatically controlled AN/SPY-1 radar, which provides target search and track and missile guidance functions. The AN/SPY-1 radar must deliver these functional capabilities under a wide range of natural environments, including rain and land and sea clutter, in the presence of electronic counter measures. Achievement of required AEGIS system effectiveness demands a high level of operational readiness.

The objective of AEGIS availability design is to produce a weapons system that is operationally ready, at an acceptable level of performance, whenever needed. Considering complexity, the AN/SPY-1 Radar System presents the greatest challenge to the achievement of this objective. Through specific examples, this paper will show how RCA approached a variety of reliability and maintainability design problems with the AN/SPY-1 radar and the overall success achieved.

The simplicity of the availability design objective can be deceptive. Many problems stand in the way and must be resolved during system design. Among the considerations that form a part of the total design approach are:

- Manning - The number of operations and maintenance personnel must be lower than existing systems with similar functions.
- Maintenance Skill Level - All on-board corrective and preventive maintenance must be within the skill capabilities of typical Navy maintenance personnel.
- On-Board Sparing - On-board sparing required to support readiness requirements must be reduced to reasonable space and cost constraints.
- Accessibility and Restoration - Access to and replacement of failed components must be rapid. This must be accomplished by use of standard tools and fixtures and available manpower and must not be a cause of damage to other components.

Thus the relatively simple availability objective translates into a number of interrelated design constraints that must be considered in a coordinated manner to assure attainment of a useful solution. At this stage of the program the first prototype equipments have been fabricated and are being tested. The material covered in this paper reflects the availability design of this prototype equipment and is presented as follows:

- Description of the principal components of the AN/SPY-1 radar as part of AEGIS
- Summary of the RCA design approach to achieve high operational availability
- Application of the design approach with examples
- Degree of accomplishment in the design for availability.

Principal Components of the AN/SPY-1 Radar

Radar Antenna Group

The AN/SPY-1 radar antenna group consists of four identical array faces. Each array face is oriented to provide radar coverage for a quadrant (90^o) such that the four faces provide full hemispherical coverage of the ship. Figure 2 pictures the rear of one of the array faces. The engineer is shown with his hand on the access door to a phase shifter driver board nest. There are 4,480 phase shifters and associated driver circuits in one array face.

The functions of the antenna are to form RF energy into a beam through phase control of its radiating elements, radiate RF energy into space, direct the beam, and receive the return signal reflected off targets in the beam path. Phase control to form and direct the beam is accomplished through phase shifter elements associated with each antenna face.

Beam Steering Control (BSC)

The AN/SPY-1 radar includes two BSC units, each serving two array faces at opposite ends of the AEGIS ship. The function of the BSC is to compute the required instructions to form and steer the antenna radar beam. The instructions are in the form of analog signals to phase shifter drivers in the radar antenna and transmitter. Figure 3

shows the BSC cabinet with the door open giving access to its components. One of the half frames is extended exposing the module nests at the top and power supplies at the bottom.

Transmitter

The AN/SPY-1 radar incorporates two transmitters, each serving two array faces at opposite ends of the AEGIS ship. The function of the transmitter is to deliver pulsed radio frequency power to two associated phased array antennas, upon receipt of timing signals and radio frequency waveforms from the signal processor. Figure 4 illustrates a final power amplifier cabinet showing access to the high power RF amplifier tubes. Each compartment can be electrically isolated to permit maintenance without radar shutdown.

Signal Processor

The AN/SPY-1 radar includes two signal processors, each serving two array faces at opposite ends of the AEGIS ship. The signal processor, operating in various modes under control of the radar control computer, generates the complex waveforms required for transmitter excitation, selects appropriate inputs from the RF receiver (part of the antenna), and processes these signals to extract detection, tracking and ECM analysis information for output to the radar control computer and the display video formatter (part of radar control).

Radar Control

The AN/SPY-1 control provides the control equipment and operator interfaces that enable radar target data to be acquired and processed in forms required by the radar and other AEGIS equipment. Radar control tasks include radar turn-on and initialization, initiation and control of target track, smoothing of track data, missile data link, formatting of radar video for display, and the analysis of radar returns.

Availability Design Approach

As already stated the AEGIS availability design objective is to produce a weapons system that is operational, at an acceptable level of performance, whenever needed. It was also stated that a good solution must fully consider related factors, including manning, maintenance skill levels, and on-board sparing. A full understanding of availability and related factors is necessary before the weapon system availability objective can be translated into specific equipment and support solutions.

To gain this understanding, a requirements analysis was initiated in the early stages of the AEGIS program to identify and relate reliability, maintainability, support and operational factors to AEGIS system effectiveness. A valuable tool used in the evolution of the analysis is the requirements structure which shows the interrelationships of system effectiveness elements. This requirements structure, see Figure 5, demonstrates that system effectiveness is not a single measure. Each system effectiveness measure would also have a unique set of functional relationships to the elements shown in Figure 5.

The first level of flow-down from system effectiveness covers:

- Performance parameters
- Operability parameters (i.e., availability and reliability)
- Command utilization.

Thus system effectiveness, as defined for the AEGIS Weapons System, can be expressed as

System Effectiveness = f (Performance, Operability, Utilization).

The requirements structure displays the interdependencies among the various requirements. Initially it provides a basis for the qualitative consideration of: (1) the significance of each requirement; (2) the order of precedence, and the interrelationships among the requirements; (3) the parameters that must become quantitative terms in one or more of the mathematical models; (4) a starting point for the allocation of subrequirements in the specification hierarchy; and (5) the definition of the interrelationship of system design disciplines (e.g., reliability, integrated logistics support and equipment design). As a convenience in relating the factors in the structure to design, four basic design objectives were defined, as follows:

- Provide the most cost-effective reliability and maintainability characteristics in the equipment building blocks. This encompasses factors at the lower tiers of the requirements structure, including access time, remove and replace times, and mean time between malfunction events (MTBE).
- Desensitize system performance to building-block malfunctions. This includes the system configuration characteristics, such as redundancy and graceful degradation, that relate system availability, reliability, MTBF and MTTR to the building blocks.
- Satisfy system requirements related to maintenance and logistics burdens. This ties the building block characteristics to system effectiveness measures associated with support factors, such as manning and sparing.
- Provide facilities for on-ship evaluation of weapons system readiness, rapid recovery to higher operability states, and control of reconfiguration alternatives. The AEGIS Operational Readiness Test System (ORTS) is at the focus of this design objective. The requirements structure shows ORTS functionally involved in fault detection, fault isolation and status reporting. A separate paper presented at the AEGIS session, "AEGIS Operational Readiness Test System (ORTS) - Design for System Effectiveness", addresses specific ORTS requirements and implementation.

Achievement of each design objective in the AN/SPY-1 radar represents close collaboration between RMA specialists, and system designers. The inherent parallelism in the AN/SPY-1 radar configuration has been exploited along

with control of building block reliability and maintainability characteristics to achieve high levels of system effectiveness at little cost. Each of the design principles is briefly discussed in the following paragraphs.

Provide the Most Cost-Effective Reliability and Maintainability Characteristics in the Equipment Building Blocks

This design objective represents an attack on the basic problems of:

- Minimizing equipment failure frequencies
- Ease of maintenance to restore operation after a malfunction. Regardless of the application of other techniques to achieve high operational availability, a prime requisite for a viable system is that the total quantity of malfunctions be kept within reasonable bounds.

The reliability and maintainability building block concept is sufficiently broad to be applicable at any level of system configuration detail. Typically the blocks fall into four categories:

- Functional entities, such as discrete parts and integrated circuits
- Packaging entities, such as module cards, line replaceable units (LRU's), throw-away modules, etc.
- Electrical interconnections and interfaces
- Mechanical interfaces.

The approach to this objective includes the process of judicious selection of a standard set of building blocks combined with a totally integrated approach to design analysis, design review, and design validation testing. The minimization of failure frequency is based on the premise that through the careful selection and application of parts and conservative circuit design it is possible to realize on-ship failure rates substantially lower than those indicated by standard prediction techniques. Specific steps that have been taken in the AN/SPY-1 reliability design include:

- Standardization to Minimize Number of Unique Types - The standardization program has succeeded in limiting the almost one million electronic parts of the AN/SPY-1 to 2,500 types of which 98 percent are standard.
- Establish Parts Derating Policy - The AEGIS Standards Manual invokes this policy on all new AEGIS designs.
- Selective Use of Established Reliability Components - Critical radar elements, such as the signal processor, have incorporated established reliability parts to enhance its reliability.
- Establish Requirements for Acceptance Testing, Screening, and Burn-in - All monolithic and hybrid integrated circuit specifications invoke the screening requirements of MIL-STD-883 Class B.

Ease of maintenance has been based on a continuous close interrelationship between the designers and maintainability engineering personnel. Specific points that have been pressed by maintainability engineering to assure ease of maintenance for AN/SPY-1 include:

- Access time - Through careful location and positioning of replaceable units access to over 95 percent of all units can be made in 10 minutes or less.
- Replacement time - Through modular design techniques most units can be removed and replaced in a few minutes. Even areas with histories of long removal and replacement times have replacement time significantly less than one hour. See, for example, Figure 4 which shows the final power amplifier tubes in the transmitter located in individual compartments for rapid access and replacement.

Desensitize System Performance to Building Block Malfunction and Repair Characteristics

When a system as complex and multi-functioned as the AN/SPY-1 is designed to reasonable cost constraints, the sheer quantity of parts constituting the system makes it essential to design to minimize malfunction effects as well as malfunction frequency. A viable system must have the inherent capability of continued operation despite failures of individual components.

The simple answer of large-scale redundancy is prohibitive from the standpoints of cost, space, and maintenance burden. RCA has chosen an approach that capitalizes on design opportunities to utilize functional modularization such that some useful level of system performance is retained despite most individual component malfunctions. This approach emphasizes features relating to both the malfunction and the subsequent repair.

Many existing systems are serial in nature, so that almost all malfunctions result in a down system. The AN/SPY-1 radar differs from this serial-type system in that most malfunctions have little or no impact on weapon system performance and almost no malfunctions have been identified that would take the radar completely down. This desensitization to malfunctions in the design has been accomplished through the application of the following techniques:

- Load Sharing - This technique involves a partitioning of performance into independent channels, so that the loss of any one channel is tolerable and permits continued useful system operation.
- Functional Modularization - By this technique alternate paths of completing a system function are kept functionally independent, that is the loss of a functional path would not result in a complete loss of the functional capability. However, the alternate path may have somewhat less effective performance for the conditions.
- Reconfiguration - This technique takes advantage of the possibility of reorganizing remaining equipment after a malfunction so as to continue operation around the failed equipment with some incremental loss in performance capability.
- Selective Redundancy - The foregoing techniques to reduce the impact of malfunctions on performance take

advantage of design opportunities with little or no increase in equipment complexity. Critical remaining serial components may be desensitized through redundancy. By this technique each candidate application of redundancy was evaluated against a criteria that relates the reliability payoff in terms of decrease in system failure rate to the added equipment complexity. With redundancy there is no loss of performance when a malfunction occurs.

Specific examples of the application of each of these techniques to the design will be given subsequently along with an indication of the reliability payoff. A figure of merit has been developed to measure this payoff called the "desensitization index". This measure is the ratio of all malfunctions to those malfunctions that would result in a "down" system:

$$\text{Desensitization Index} = \frac{\text{Malfunction Rate (all malfunctions)}}{\text{System Failure Rate}}.$$

The degree of success in applying these techniques to the reliability design of the AN/SPY-1 is a desensitization index of 25. This means only one malfunction in 25 will result in system failure. It is noted here that the desensitization index depends on the definition of satisfactory operational performance. The numerical indices given in this paper apply to essentially full operational performance. If more relaxed but still significant operational performance was selected as a criteria the index for AN/SPY-1 would be considerably higher than the 25 cited above.

Further benefits to be derived from desensitization relate to repair actions for restoring the system to full performance. For those malfunctions that do not disable the system, it is essential that consideration be given to:

- Provide the capability of fault detection, fault isolation, access, removal, replacement and verification without shutdown of operating functions (on-line maintenance)
- To allow for planned deferral of maintenance without significant compromise of system reliability characteristics or extended periods of low weapon system performance.

Both of these techniques have been provided in the reliability design of the AN/SPY-1 and are discussed in the following paragraphs.

<u>Satisfy System Requirements Relating to Maintenance and Logistic Burdens</u>

Specific techniques for implementing this design objective are closely interrelated to the reliability design for R&M building block characteristics and the desensitization to malfunctions. This objective has been identified to focus attention on the essential nature of system support in the total availability picture. Perhaps the key payoff factor that can be identified is through the exercise of maintenance deferral options.

For most situations where the design incorporates desensitization to malfunction characteristics, maintenance deferral can be used to achieve a more effective utilization of maintenance personnel. Where the incremental degradation characteristics achieved through desensitization are sufficiently small, it is possible to defer to dockside operations thereby:

- Minimizing or eliminating the need to provide shipboard sparing of the affected items.
- Reducing at-sea maintenance manning requirements.

Major savings in manning and sparing have been realized through the application of these techniques to the AN/SPY-1 radar antenna. This treatment of the antenna reliability design will be discussed in greater detail further on in this paper in a discussion of availability design implementation.

<u>Provide Facilities for Aboard-Ship Evaluation of Weapons System Readiness, Rapid Recovery to Higher Operability States, and Control of Configuration Alternatives</u>

The ability to fully exploit the intrinsic availability potential of the AN/SPY-1 is dependent upon having accurate knowledge at all times of the:

- True condition of the system
- Configuration options available during maintenance or casualty modes
- Performance capability of each configuration alternative
- Options and penalties associated with shutting down for maintenance, performing on-line maintenance, or of deferral of maintenance actions.

The system design must include readiness measurement and evaluation machinery that provides all levels of shipboard operational and command personnel with a continuing evaluation of system operational readiness. The basis for the selection of operational procedures, configuration alternatives, and problem identification for specific situations should be an integral part of the system design.

This fourth design objective is concerned with a factor that, although seldom treated explicity or quantitatively, is best designed-in during system synthesis. This factor is the capability to fully realize the equipments' inherent operability characteristics in the shipboard environment. RCA's solution for AEGIS is to provide a fully integrated, system-level monitoring and operability testing facility called the Operational Readiness Test System (ORTS). ORTS consists of an organization of equipment hardware, computer programs, shipboard procedural controls over resources and appropriate documentation that provide:

- Knowledge of system status
- Fault identification
- Available configuration alternatives.

ORTS parameters have been mathematically modeled and exercised on computer programs. A set of parameter values has been selected for the AN/SPY-1 radar as requirements to retain the inherent availability characteristics of the design.

A more detailed discussion of ORTS and its relation to availability is contained in another of this set of AEGIS papers, "AEGIS Operational Readiness Test System (ORTS) - Design for System Effectiveness". This paper addresses:

- The relation between system effectiveness and the ORTS function
- The development of ORTS requirements
- Key system implementation aspects of ORTS.

Availability Design Implementation

This portion of the paper addresses the most critical availability design problems encountered in the AN/SPY-1 radar and their solutions. In each case a measure of payoff is provided to indicate the relative impact of the solution. Thus, for example, design solutions that take advantage of techniques, such as load sharing and redundancy, are evaluated by their desensitization index, while design solutions based on parts improvements are evaluated by a reliability improvement ratio. The design solutions discussed have been incorporated in the equipments now being readied for test. Additional potential design modifications have been identified for further availability improvement through continued engineering evaluation. These further improvements are being evaluated jointly by RCA and the Navy for the future.

Antenna Beam Steering

The most significant area in the AN/SPY-1 from a parts complexity standpoint is antenna beam steering. Approximately 50 percent of the total parts population of the AN/SPY-1 are included in the antenna transmit/receive phase shifters and their associated drivers. There are approximately 20, 000 antenna phase shifters and driver circuits divided among the four antenna array faces. For a successful reliability design, this area must have a complete solution.

An obvious approach to this solution is through load sharing. Fortunately, it was an easy task to design the phase shifter paths to be essentially independent of each other, so that any single phase shifter or driver failure would have a negligible impact on performance. Through this design for circuit independence the impact of a malfunction on radar detection range is less than 2/100 of 1 percent. This is a good start but is a long way from the final solution.

One factor that must be listed as a major problem is the location of the phase shifter in the antenna waveguide. One look at the rear of the antenna, see Figure 2, indicates the impossible task of having these phase shifters in a readily accessible position. Major dissassembly of the antenna is required. However, the small incremental performance degradation for each phase shifter malfunction points the way to a solution. If the unit malfunction rate can be held sufficiently low, it is possible to defer all phase shifter maintenance to in-port maintenance.

Performance studies of array degradation have shown that up to 10 percent of the array elements can be lost through phase shifter or driver failures without significant degradation in weapon system performance. Further, by placing emphasis on phaser reliability it is felt that through scr ning and burn-in techniques failure rates in the order of 0.5 failures per million hours can be achieved based on experience in other programs.

The question of what the phaser reliability should be, in order to assure that no shipboard maintenance is required for the antenna phasers, has been treated in a parametric study. The results of this study are shown in Figure 6 for an 18 month maintenance cycle and a 10 percent array degradation criteria. From these results it is concluded that if the phaser failure rate is held below six failures per million hours the full 18 months maintenance deferral period would pass with a very high probability of performance exceeding the degradation criteria. Based on these results there is high confidence in the feasibility of this maintenance policy.

The phase shifter drivers are a much easier problem from a maintainability standpoint and they have been located in accessible nests at the rear of the antenna array. However, by their number the potential maintenance load is substantial. Parametric studies similar to those made on the phase shifter indicated that maintenance deferral to each in-port period is quite feasible with very little performance impact.

By taking full advantage of this reliability design approach and exploiting its effects through the maintenance policy the following reliability/maintainability payoff was achieved:

- Desensitization - Weapon system performance has been almost completely desensitized to the malfunction of antenna beam forming components.
- Sparing - It is not necessary to carry on-board spares for the antenna phase shifters or their associated driver boards.
- Manning - Elimination of the need for at-sea maintenance of antenna phase shifters and driver boards reduces maintenance loading for the radar system by approximately 30 percent.

Transmitter

After the antenna beam steering components, the AN/SPY-1 transmitter presents the next highest malfunction rate in the radar. Here again a good design solution is required to achieve the desired levels of availability. Three basic approaches to reliability design were selected for the transmitter: (1) redundancy in the lower power stages of the transmitter, (2) load sharing in the final power amplifier, and (3) on-line maintenance.

The transmitter, see Figure 7, is divided into three major stages:

- Input amplifier stage
- Pre-driver/driver stage
- Final power amplifier stage.

The input amplifier stage consists of two identical low power TWT amplifiers, one on-line and the other in a powered standby state feeding a dummy load. Should the

on-line amplifier fail, it is switched off-line and the standby unit switched on-line with immediate resumption of operation. Repairs to either unit may be made while in the off-line position.

The pre-driver/driver stage provides the power necessary to drive the 32 power amplifiers that provide RF power to each of two array faces. This stage consists of 4 cabinets, each containing two TWT-CFA channels. In operation three of four cabinets are required, with the fourth in ready. Should any one of the four cabinets fail, it is switched off-line, and maintenance may be performed to restore its operation without interfering with transmitter operation. The switching action provides for rebalancing of the remaining three units to provide optimum power combining.

The final amplifier stage, which provides the pulsed RF power to the array faces, is made up of 64 separate power amplifier units, 32 per face. These are CFA amplifiers arranged in groups of four. Two groups of amplifiers, four amplifiers from each face, share a common power supply (including modulators). The final amplifier stage includes eight such common power supplies. By this arrangement the malfunction effects of CFA's, modulators, or high voltage power supplies are limited to the RF power associated with the failed item. For example, a modulator failure would result in the loss of power from 4 CFA's (1/8 of total transmitted power) in one face. The impact on AEGIS performance is less than 7 percent loss in detection range capability, which is small. Each CFA amplifier is contained in a drawer that slides open for maintenance (see Figure 4). Each amplifier can be isolated from its power supply and other amplifiers (via disconnects and RF shutters in the waveguide) so that it can be repaired without interfering with operation of the remaining units. Also, each power supply is independent and can be maintained without interference to the others.

Although the design has excellent reliability and on-line maintenance characteristics, further improvement was possible relative to transmitter sparing, maintenance manning, and life cycle cost. At this writing specific design actions have been taken or are under evaluation to simplify the transmitter without a significant effect on its reliability characteristics. These design actions include:

- Elimination of 1/2 of the final amplifier CFA's. Incorporation of an RF switch at the output of each remaining CFA to switch its power to either array face.
- Elimination of eight high voltage power supplies (HVPS's) between the transmitters at either end of the ship and the sharing of the remaining 8 HVPS's between both transmitters.

Detailed evaluations indicate that these significant design simplifications can be accomplished without decreasing transmitter reliability characteristics. Current plans are to incorporate the new configuration as a modification to the engineering development model.

As a measure of reliability design accomplishment the desensitization index for the transmitter is approximately 100.

Power Supplies and Cooling

The next ranking reliability design problems that required solution were the essential supporting functions of low voltage power and cooling. This was identified as a prime area for the application of selective redundancy. To this end the AN/SPY-1 reliability design incorporates the following redundancy applications:

Low Voltage Power

- Essentially all power supplies in electronic cabinets
- Antenna RF receiver power supplies
- Phase Shifter driver power supplies

Cooling

- All cabinet cooling fans
- Antenna air cooling fans.

After factoring in the design solutions already indicated for the antenna beam steering and transmitter, this step approximately doubles the AN/SPY-1 system reliability.

This leaves us with the most difficult reliability problem in the AN/SPY-1 radar, the signal processor.

Signal Processor

The signal processor is the most functionally complex unit in the AEGIS system. The unit is complex because it must be capable of handling multiple modes and functions simultaneously. A total combination of eleven modes/submodes and functions are provided by the signal processor: search-in-clear, track-in-clear, moving target indicator (MTI) search, MTI track, burn through; passive search and track, cover pulse, barrage jamming detection, prelook, missile communications and target definition. Much of the unit is channelized on a frequency basis, resulting in identical hardware being used to process four frequency bands simultaneously. Further, in some modes both in-phased and quadrature components of each frequency are processed. To fulfill these functions the signal processor equipment has been divided into seven cabinets; I/O buffer-synchronizer, A/D converter, waveform generator, common IF processor, MTI mainlobe processor, MTI sidelobe processor and the pulse compression processor. The simplified interface between the seven cabinets of the signal processor is shown in Figure 8.

The design for reliability of the signal processor incorporates selected redundancy, integrated circuit screening and utilization of the incremental degradation possibilities of the frequency channelization. Through these reliability design measures a desensitization index of 4 has been achieved in the signal processor. Following are key reliability design features of the signal processor.

The signal processor is the dominant reliability series link in the AN/SPY-1 Radar System and in the AEGIS Weapons System. As such it has received a sharp focus of

attention throughout the program and will continue to receive this attention as long as it remains a dominant factor. Major steps already taken in design for reliability will be reviewed, followed by a review of the results of recent reliability studies, which are being evaluated as potential design modifications.

A major step toward improving reliability of the signal processor at the parts level (i.e., building block level) was the inclusion of the screening requirements of MIL STD 883 on the procurement specifications of all monolithic and hybrid I.C.'s. Based on MIL HDBK 217 failure rates, this step leads to an initial parts reliability improvement for the signal processor of up to 4 to 1.

It was recognized at an early point in design that approximately 50 percent of the total parts failure rate would be attributable to power supplies and cabinet cooling fans. Based on this a design decision was made to incorporate redundant configurations for all cabinet cooling fans and high current low voltage power supplies. These decisions resulted in more than a 2-to-1 desensitization improvement. The remainder of the desensitization factor of four is accounted for by the frequency channelization and other part failures that have minor effect on weapon system performance.

Although a desensitization index of four is quite good, both RCA and the Navy felt that further significant improvement could be achieved. If all of the recommended changes are eventually included in the final design, a total reliability improvement of ten would be realized as contrasted to the current desensitization factor of four. Recommended potential design modifications are as follows:

1. Central Signal Processor - Replace the fore and aft signal processors with a single central signal processor. This leads to a 60 to 70 percent reliability improvement plus logistic and manning benefits.

2. Adaptive Frequency Channels - Use ORTS to recognize a failed frequency channel and exclude any data in the failed channel from further processing. This would eliminate the processing of noise in a failed frequency channel and the consequent loss in accuracy of angle error computations.

3. Increased Frequency Channelization - Through a redesign of the A/D converter arrangement and associated circuitry it is possible to limit the effect of a single failure to two-frequency channels rather than all four frequency channels. In addition through the sensing of the failure by ORTS and a change to a two-frequency waveform, the failure effect on performance would be negligible.

4. Selected Redundancy - Through a reorganization of some components and the addition of others selective redundancy can be applied to significant reliability series links in the synchronizer and waveform generator functions.

The reliability changes in items 2, 3, and 4 would lead to a 100 percent reliability improvement. The desensitization index for a central signal processor with the changes indicated in 2, 3 and 4 would be approximately eight.

Radar Control

The final major area that remains to be resolved is the radar control computers. Computer program design development to this point has been focused on integrating and solving the very complex AN/SPY-1 control problems. Still to be addressed are the computer reconfiguration options in case of computer failure. After successful demonstration of the AN/SPY-1 control programs this subject will be placed in the spotlight. This step is vital since the AEGIS computer complex without reconfiguration would be the only remaining major series link whose failure would result in a down weapon system.

Summary of AN/SPY-1 RMA Design Achievements

The RMA program for AEGIS has had significant success in achievement of the availability design objectives for the AN/SPY-1 Radar System. A measure of this success is the design availability of the AN/SPY-1 radar:

- 0.99 for full operational performance
- 0.995 for degraded but useful operational performance.

Further, these results have been achieved while radar support requirements have been controlled within bounds. Support requirements include manning, maintenance skill levels, sparing, accessibility and restoration.

Radar availability design characteristics that contribute to this achievement are summarized as follows:

Availability

- Part reliability was controlled and improved through standardization, enforced derating policy, selected screening requirements, and selected use of established reliability parts.
- High MTBF was achieved through a vigorous exploitation of design opportunities to desensitize performance to malfunctions. The result is that only one in 25 malfunctions has measurable impact on operational performance. Essentially no malfunctions take the system completely down.
- Low maintenance times are the result of rapid fault detection and isolation by ORTS, modular design, and access to most replaceable components in less than 10 minutes.
- Full and accurate fault detection coverage by ORTS protects against degradation below inherent availability capability.
- On-line maintenance capability for most desensitized components permits complete corrective maintenance without taking the radar down.

Maintenance Manning

- Stringent control of parts reliability and equipment maintainability characteristics has had a direct effect in lowering maintenance burden.

- Deferral of antenna phase shifter and driver corrective maintenance to in-port period has reduced total maintenance man hours by 30 percent.

- Short term deferral of corrective maintenance of other desensitized components, such as low voltage power supplies, permits a more effective utilization of manpower.

Maintenance Skill Level

- ORTS provides rapid automatic fault detection and isolation to the replaceable unit. The unambiguous ORTS indications simplifies the maintenance task as well as associated maintenance documentation so as to be well within the normal Navy maintenance skill capabilities.

- All replaceable units are modules, assemblies, or major parts (e.g., transmitter tube) that require minimal mechanical skills for removal and replacement.

On-Board Sparing

- The desensitization of performance to malfunctions significantly reduces on-board sparing to a relatively small set of critical components. In effect, the desensitization characteristic permits the system to live off itself between in-port periods.

Accessibility

- Through design care in locating replaceable components and in the mechanical design of enclosures, access can be made to 95 percent of all replaceable components in 10 minutes or less.

Restoration

- Replaceable component weight has been held to 40 pounds or less for all but a few special items (e.g., transmitter CFA final power amplifier).

- Modular design throughout the radar contributes to rapid removal and replacement.

Acknowledgements

The design of the AN/SPY-1 Radar System is the culminations of the efforts of many contributors not only at RCA but by members of the staff at APL, NAVSEC, and NAVORD. In particular, the author wishes to recognize the Navy AN/SPY-1 Project Director, R. Hill of NAVSEC and the RCA AN/SPY-1 Project Manager, Mr. H. Grossman.

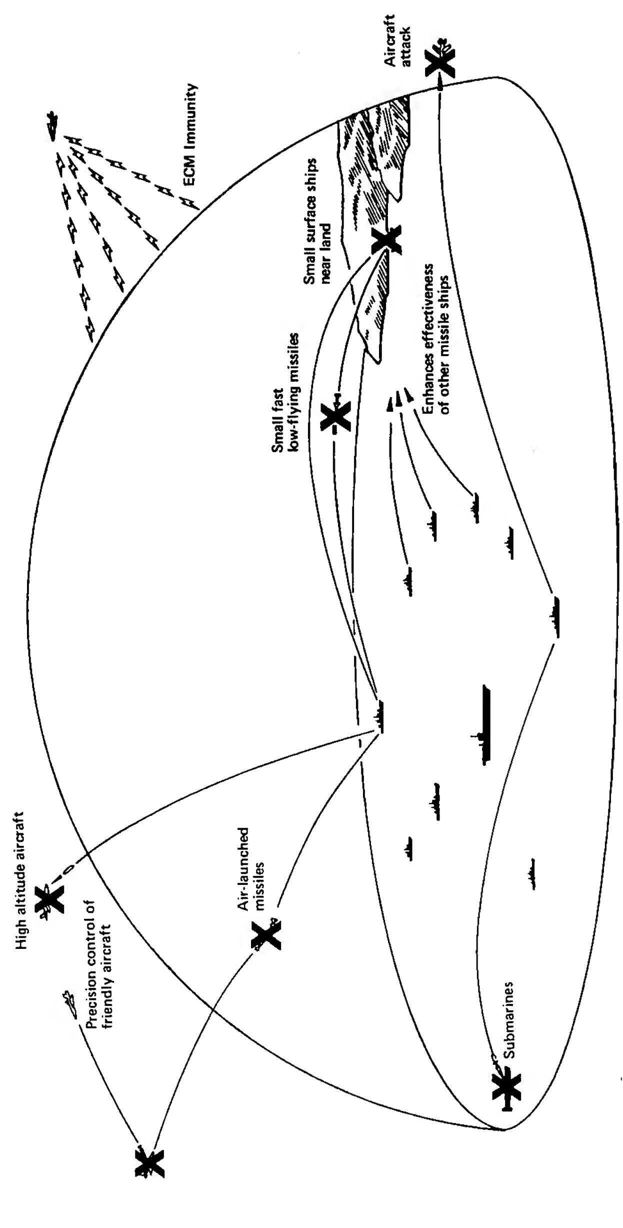


Figure 1. AEGIS Escorts Engaging a Variety of Targets

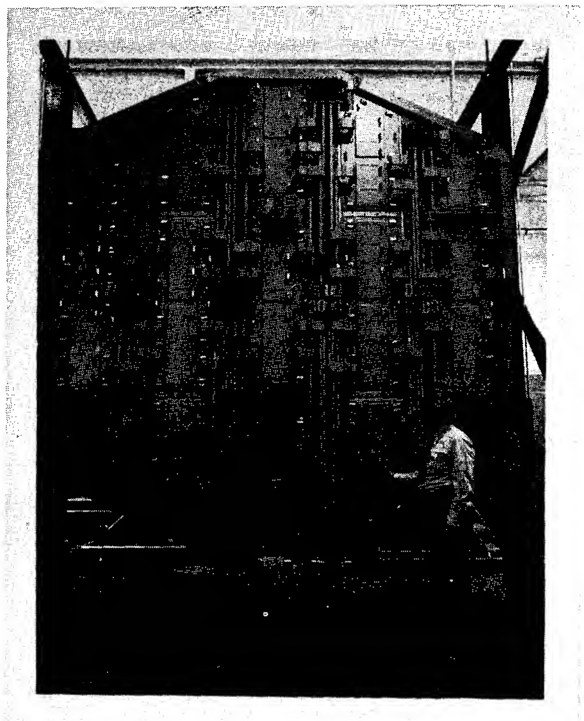


Figure 2. Antenna Array for AN/SPY-1 Radar (Back View)

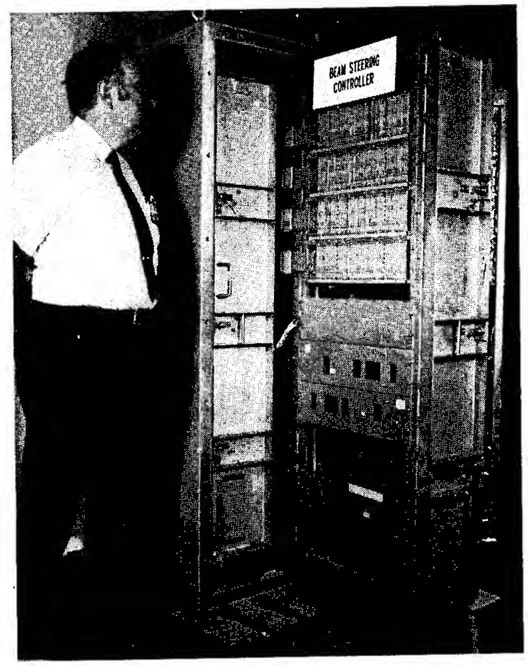


Figure 3. Beam Steering Controller for AN/SPY-1 Radar

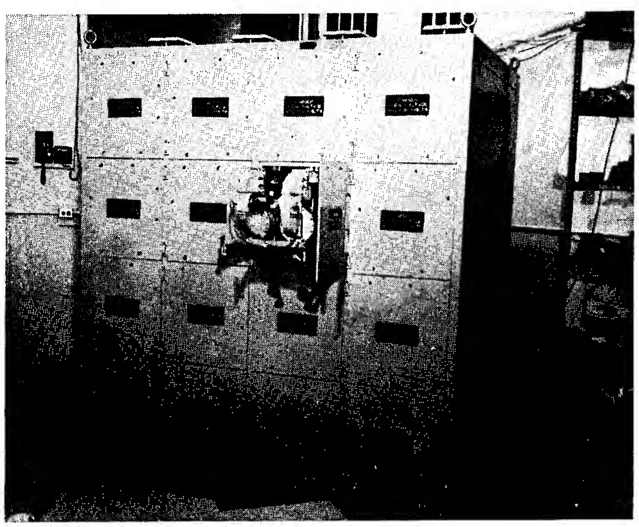


Figure 4. Transmitter Final Power Amplifier Cabinet for AN/SPY-1 Radar

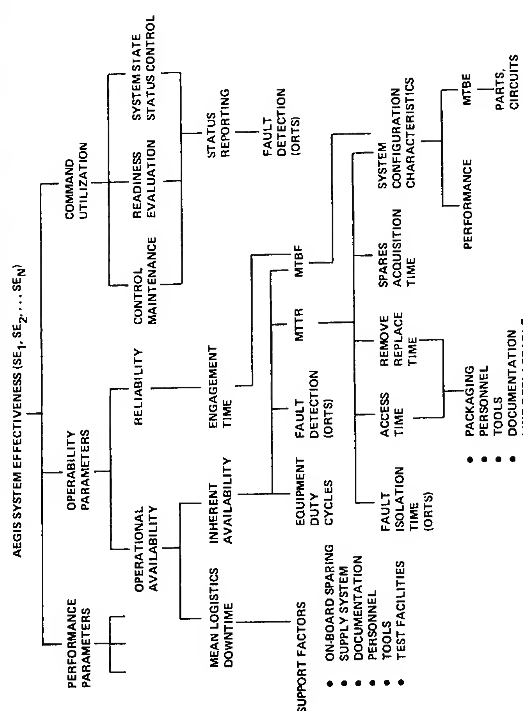


Figure 5. AEGIS Requirements Structure

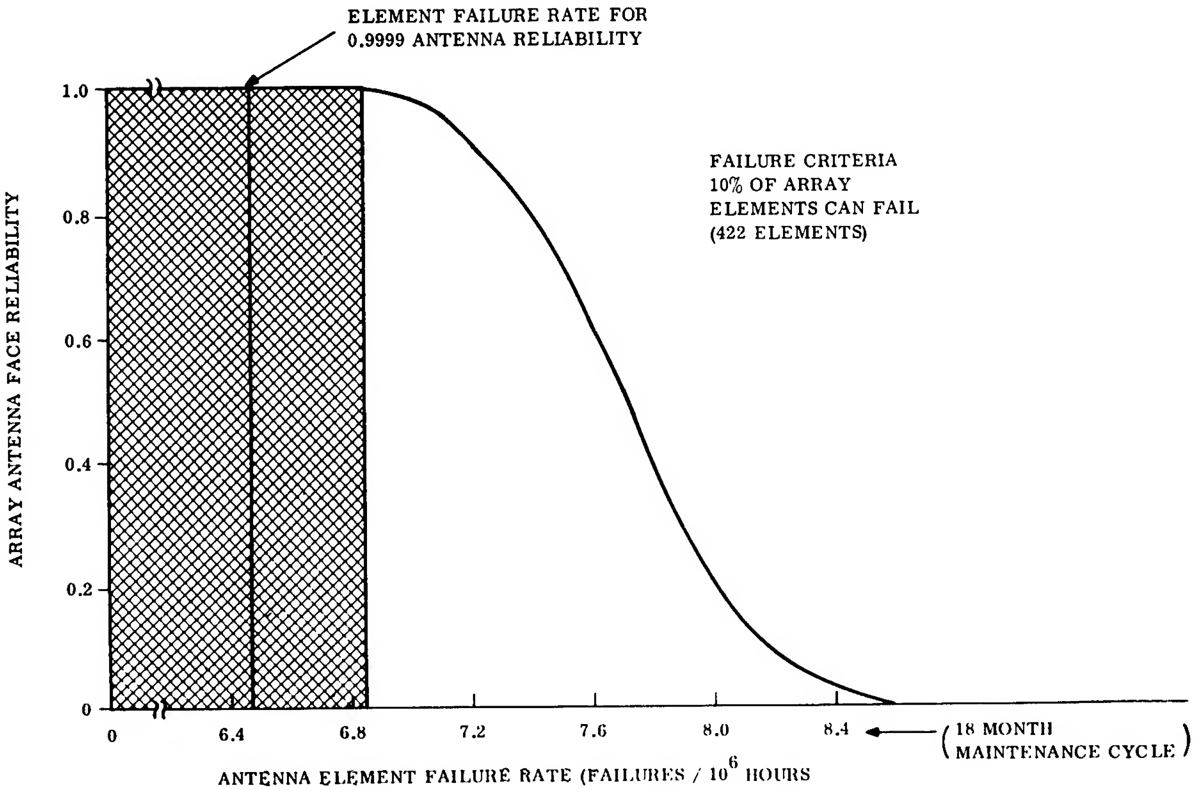


Figure 6. Antenna Reliability for No Repair Maintenance or Overhaul Cycle versus Element Failure Rate

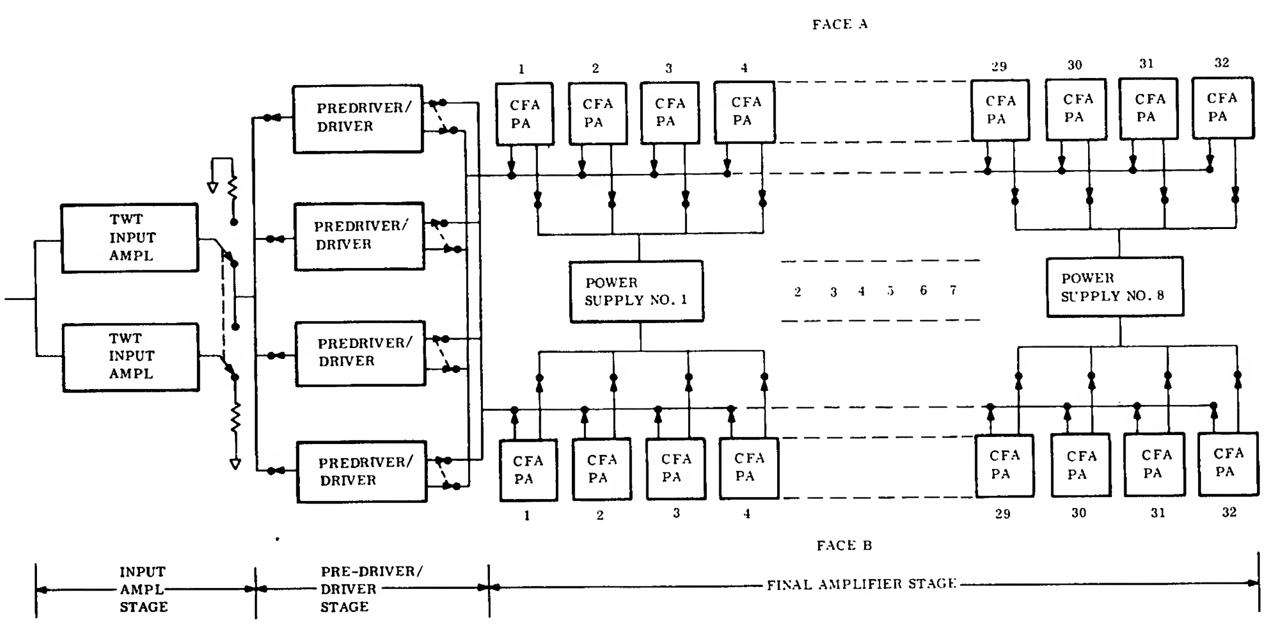


Figure 7. AN/SPY-1 Transmitter Functional Block Diagram

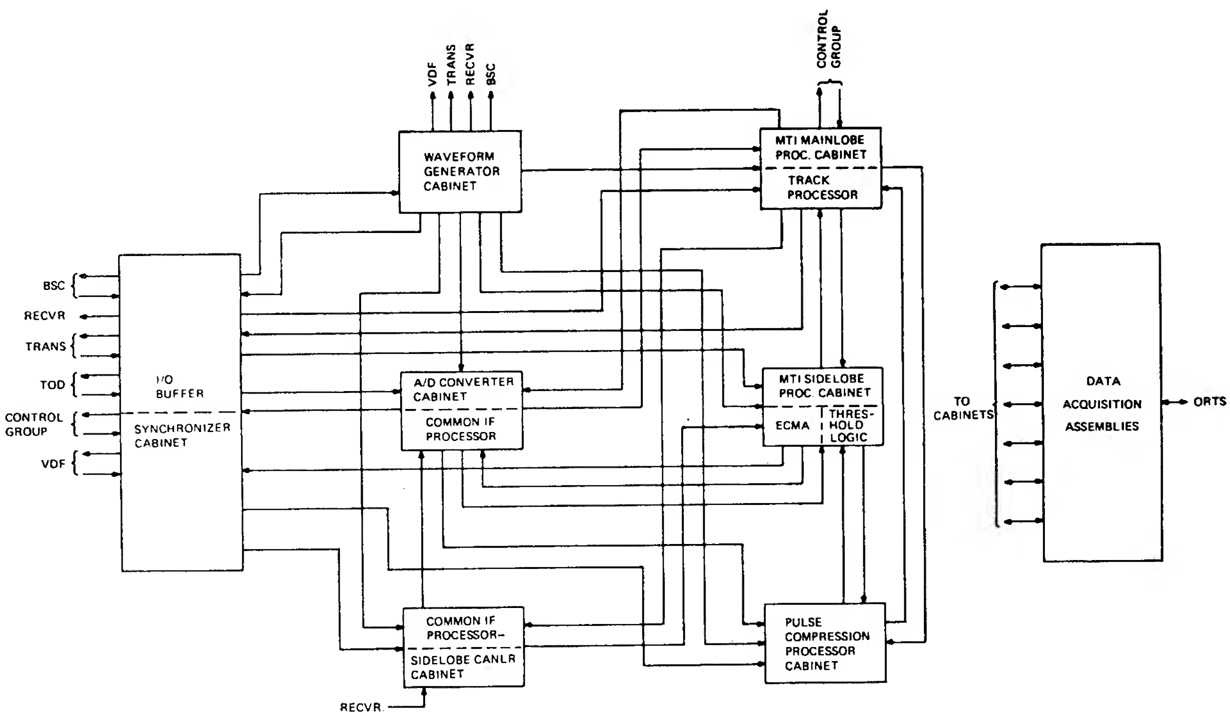


Figure 8. Signal Processor Functional Block Diagram

AEGIS DEMINERALIZER/WATER COOLER - DESIGN FOR AVAILABILITY

S. R. Gladstone
RCA Missile and Surface Radar Division
Moorestown, New Jersey

INDEX SERIAL NUMBER – 1047

Summary

The AEGIS Weapons System is an advanced shipboard missile system characterized by the fast reaction and unprecedented fire power required to defend the fleet against the sophisticated airborne and surface threats of 1975 and beyond. To meet the specified AEGIS operational readiness requirements it is vital that the shipboard cooling system be continually available to support the complex electronic systems that constitute the weapons system. Navy experience has revealed that cooling systems are a frequent cause of inoperative shipboard weapon systems. This paper cites some of the major cooling system problems that have been encountered by the Navy. These problems are treated in the context of the design for availability of the AEGIS Demineralizer/Water Cooler. The availability design approach that is described is directed toward the elimination of historical problems and the inclusion of the positive design features that will ensure high cooling system availability. Those features that are incorporated in the design as a result of this design process are reviewed.

Introduction

The AEGIS Weapons System currently being developed for the U.S. Navy is an advanced shipboard missile system characterized by the rapid reaction and high fire power required to defend the fleet against hostile aircraft, missile, and surface threats of 1975 and beyond. Seven major systems constitute the AEGIS Weapons System: the AN/SPY-1 Radar System, Command and Control, Weapon Direction System, Fire Control System, Operational Readiness Test System, Guided Missile Launching System, and the Missile System. Although each of these systems has a rigidly specified function, the functions interact strongly when the systems are integrated and in the operational mode. To ensure that all systems operate together as an entity and meet specified AEGIS operational readiness requirements, it is critically important that the shipboard cooling system be continually available to support these systems. This paper will address the problems and the resulting solutions associated with the design and development of the AEGIS Demineralized Water Cooling System, depicted in gray in Figure 1. The interrelationships between the various shipboard cooling services are also depicted in Figure 1, which shows both the air and water cooling systems and the compartments of the ship serviced by the systems. Also shown are the air and water cooling systems that are dedicated to AEGIS.

From the inception of this program the Navy has indicated that a high priority was to be assigned to cooling. Experience at sea had identified cooling systems as being a frequent cause for inoperative weapon systems. Fortunately, this same experience revealed several particularly troublesome aspects of these cooling systems. During the development of AEGIS every effort was expended to solve these problems. The cooling system that has resulted is a weight-effective combination of air and water-cooling equipment designed to give the AEGIS Weapons System the same high operational availability as the ship's prime power, steering, and propulsion.

It is the purpose of this paper to:

- Provide a historic background on existing shipboard cooling systems to illustrate the problem areas associated with such systems.
- Summarize the performance and Reliability and Maintainability (R&M) requirements allocated to the Demineralized Water Cooling System.
- Illustrate the role that the R&M discipline had in influencing the system engineering and design process.
- Provide a brief functional description of the Demineralized Water Cooling System.
- Summarize the significant features incorporated in the design.

Historic Background

Historically, difficulties with the ship's cooling services have been a significant source of system downtime when used to support large-scale auxiliary electronic systems. One reason for this is that, in the past, electronic systems have been simply "plugged in" to already-existing ship's cooling. First-hand examination of shipboard systems has revealed other difficulties, including poor component performance, lack of redundancy, and a design that does not allow for corrective maintenance without shutdown. An examination of experience data concerning existing shipboard water cooling systems has revealed the following critical problem areas.

- Turbulence Erosion: Erosion within heat exchangers has been determined to be a cause of some Navy cooling system failures. This problem was identified by Naval Ship Missile Systems Engineering Station (NSMSES) personnel as being due to the extreme turbulence of salt water at the input of these units.
- Demineralized Water Contamination: Leakage of salt water into demineralized water within heat exchangers has been reported by NSMSES and NAVSEC. This has occurred at the seal between the tubes and tube sheets.
- Electrolytic Corrosion: This type of corrosion is caused by the use of dissimilar metals in contact with a common electrolyte. The use of sea water and the presence of differential electrical potentials accelerates the corrosion process and produces a particularly difficult problem in electronic cooling systems.
- Low Reliability: Existing shipboard water-cooling systems utilize a single-thread, or serial, reliability approach. Thus, any failure necessitates the shutdown of the water-cooling equipment and the radar system serviced by that cooling system.
- Inadequate Maintainability: Most shipboard water-cooling systems used for electronic equipment evidence the fact that they include equipment

"patched" into already-existing ship's cooling. This added cooling equipment is often placed into compartment spaces in configurations that make maintenance impossible and where leaks and/or condensation can cause electrical failure.

AEGIS Demineralized Water System

Figure 2 shows a simplified block diagram of the Demineralized Water Cooling System. The system consists of a MARK 1 MOD 0 Demineralizer/Water Cooler, associated cooling loops, instruments, and controls for cooling the AN/SPY-1 Radar and the MARK 99 Mod 0 Fire Control Systems.* The system is effectively divided into two identical systems, one dedicated to each Weapon System Housing, and performs the functions of purification, cooling and pumping.

Requirements Analysis

The performance requirements allocated to these functions were used to determine the gross types and sizes of equipments utilized in the design approach. The allocated performance requirements are as follows:

- Purification: Demineralize and purify cooling water to the following requirements: 0.5 ppm maximum O_2, 0.5 micron maximum particulate size, and 2.0 micromhos/cm maximum conductivity. Although the water supplied to this system may be distilled and therefore relatively pure, contaminants are picked up as the water is circulated in the cooling cycle. The removal of these impurities is a major function of the AEGIS demineralized water system.

- Cooling: Supply demineralized water to the AN/SPY-1 Radar and MARK 99 Fire Control Systems at a temperature of 86°F to 105°F, dissipating a maximum heat load of 1,038,464 Btu/hr.

- Pumping: Supply demineralized cooling water to the electronics at a maximum flow rate of 203 gpm, a maximum hydrostatic pressure of 150 psig, and a minimum pressure differential across each item of unit-level equipment (i.e., cabinet or console) of 70 psig (an exception to MIL-E-16400 imposed by NAVSEC and state-of-the-art high-power tube requirements).

The final heat sink for the AEGIS demineralized water system is ship's sea water, which is specified in NAVSHIPS 0902-019-4000 as having a minimum temperature of 28°F. A minimum sea water temperature of 28°F and a minimum operating temperature for AEGIS demineralized water of 86°F implies a water warmup requirement, and this capability has been incorporated into the design.

The R&M requirement allocations for the MARK 1 Mod 0 Demineralizer/Water Cooler were set at an MTBF of 2500 hours and an MTTR of 2 hours. These numerical requirements, graceful degradation requirements, and the elimination of R&M design problems are key factors in determining the quantities and specific types of equipments. Satisfaction of these requirements, and the elimination of R&M design problems, will be discussed later in this paper.

*The AN/SPY-1 Radar System is an electronically scanning multi-function array radar that is utilized for the detection and tracking of targets. The MARK 99 MOD 0 Fire Control System contains the guidance illumination radars for the missile portion of the Weapons System.

Design Process

The design for the Demineralizer/Water Cooler was essentially a two-step process: a preliminary design and a final design. The preliminary design effort was directed at meeting the performance requirements in terms of the necessary generic types of equipments. At this time consideration was also given to eliminating the problem areas that exist in todays shipboard water cooling systems, including those mentioned in the historical background. The final design process began when the preliminary design was assessed for the R&M characteristics. This phase of the design concerned factors such as cooling system component and material selection, selective redundancy, and the provisions for fault detection and isolation.

The modified design that evolved from this effort was then subjected to various design reviews with the Navy and their consultants, which resulted in further minor design modifications. The final design, shown in Figure 3, will be subjected to verification testing at the LBTS (Land Based Test Site) at Moorestown, N.J., during the AEGIS engineering development model integration at the end of 1972. Additional verification testing will be conducted in USS NORTON SOUND during the AEGIS engineering development model evaluation in 1973. A functional description of the MARK 1 MOD 0 Demineralizer/Water Cooler (Figure 3) operation is presented in a later portion of this paper.

R&M Assessment. During the R&M assessment of the preliminary design, the initial problem involved obtaining meaningful and representative failure-rate data for the mechanical and electromechanical items that comprised the majority of the equipment. The normal sources for failure rate data, i.e., MIL-HDBK-217 and FARADA, were of little use, and component manufacturers were unable to provide any quantitive data. However, two sources were found that contained quantitative data on representative equipments:

- "State of the Art Assessment of R&M as Applied to Ships Systems" Proceedings, 1969 Annual Symposium on Reliability, Pages 133-145 (incl.).

- "Reliability Physics (The Physics of Failure)" Proceedings, Ninth National Symposium on Reliability and Quality Control, Pages 43-57 (incl.).

Utilizing these sources as a data base and selecting the upper limit of the failure rate distribution, failure rates were modified to reflect the relative severity of the usage environment. These modifications reflected an increase in the failure rate for the valves and heat exchangers that were in the sea water cooling loop. For those items where failure rates could not be located, engineering judgment was used to determine a suitable rate based on similarity to existing items with known failure rates.

The results of the Failure Modes and Effects Analysis (FMEA) conducted on the preliminary design indicated that the reliability requirement could not be achieved without the application of selective redundancy. In addition, the results also indicated that the pump assembly and temperature control valve assembly were the major contributors to the total failure rate.

Pump Assembly. The initial design approach was to consider one on-line pump assembly plus a standby. The concept was that, upon detection of low pressure from the on-line pump, the standby pump would be automatically started and cut into the load. This approach did not fulfill the conditions of redundancy

and was discarded when an analysis revealed that the pressure buildup on the standby pump was too slow. The impact of the slow pressure buildup was that the flow rate could not be maintained at a sufficiently high rate to preclude "drop-out" of equipment flow switches prior to the cut-in of the standby pump. The final design approach is to have two on-line pumps sharing the load. Each pump is capable of maintaining sufficient pressure to ensure proper flow rates. In the case of failure of either pump, the remaining pump is adequate to continue operation of the demineralizer/watercooler without degradation of performance. Because the defective pump is isolated from the line by "stop-and check" valves, the design approach fulfills the requirements for a redundant configuration.

Temperature Control Valve. The results of the FMEA indicated that a redundant configuration was also necessary for the thermostatic control valve assembly. The part with the highest failure rate for the thermostatic control valve is the bellows or thermal element. When this element fails, the design of the valve is such that the valve "bypasses" and allows 100% of the water at the high temperature inlet to flow through the outlet. Two thermostatic control valves are arranged so that the high-temperature inlet of the second valve is connected to the outlet of the first valve, and cooled water is applied to the low-temperature inlet of each valve. Thus, failure of either valve will not result in an out-of-tolerance water temperature and the redundant configuration is achieved.

Other Design Improvements. Experience with existing ship's cooling systems has shown that sea-water-to-demineralized-water heat exchanger failures and impaired heat exchange characteristics have been caused by:

Turbulence erosion

Cavitation

Sea water contamination (sea weed)

Sea water to demineralized water leaks.

During the design process the turbulence erosion and cavitation problems were solved by locating the pressure reducing orifices at least 15 pipe diameters upstream from the heat exchangers. This eliminates the highly turbulent flow in the heat exchanger bonnets that has caused local erosion and reduced heat exchange characteristics.

Sea water contamination has been eliminated by use of a duplex basket strainer in the supply line. This unit contains two separate strainers with only one required at any one time. The basket mesh openings are smaller than the heat exchanger tubes, which reduces the probability of sea water contamination in the heat exchangers.

An in-place, spare, heat exchanger has been provided in the design, which can be put on-line by means of manually operated valves. The purpose of the spare heat exchanger is to eliminate AEGIS Weapons System downtime during maintenance or failure of the first heat exchanger.

Contamination of demineralized water by sea water is caused by leakage at the tube/tube-sheet connection in the heat exchanger. To eliminate this problem double tube-sheet heat exchangers are utilized in the design. As shown in Figure 4, this construction provides a void between the tube sheets that acts as a sea water drain if leakage occurs.

Contamination of the demineralized water is greatly reduced by the utilization of corrosion-resistant materials wherever possible for components and piping in contact with demineralized water. The use of brass is restricted to only those components not available in corrosion resistant materials. Where brass is utilized, the zinc content is limited to a nominal 15%, such as QQ-B-626 or QQ-B-613.

Results of R&M Predictive Assessment. The R&M predictive assessment of the final design indicated an MTBF in excess of 6500 hours and an MTTR of 1.5 hours. These predicted parameters indicate an adequate design margin and minimize the risk in fulfillment of the MTBF requirement of 2500 hours and the MTTR requirement of 2 hours.

Functional Description of MARK 1 MOD 0 Demineralizer/Watercooler

Referring to Figure 3, and starting at the pumps, cooling water flows through a flowmeter and is split into two streams, one flowing through the heat exchanger in use and the other through a bypass line. These streams are mixed in the temperature regulating valves, which control water temperature within a range of 86°F to 105°F (the actual temperature is dependent upon the valve setting, heat input, and sea water temperature). Downstream of the temperature regulating valves high and low-temperature switches sense abnormal temperatures and send signals to a local alarm panel. A 60-mesh duplex in-line strainer is located in the output line. As the line leaves the electronic cooling equipment room it splits into two lines, each containing a motor-driven shutoff valve. A low-flow switch is located in each return line ahead of the expansion tank. Air ejectors are located at all high points in the piping for automatic air bleeding.

Cooling water is piped to the deckhouse through two separate lines, one to the AN/SPY-1 Radar System and the other to the MARK 99 MOD 0 Fire Control System, as shown in Figure 2. If there is a ruptured supply or return line, only its associated segment will be shut down. In the case of a break, a flow switch in the damaged system's return line will automatically close a motor driven valve located in the supply line in the electronic cooling equipment room, thus preventing the total loss of demineralized water. This flow switch will also activate a local alarm and send a fault signal to ORTS. A check valve in the return line at the same location will prevent system drainage through the rupture.

The pump head of 130 psi is sufficient to supply a pressure differential at the electronic cabinets of 70 psi, or more. Ship's sea water is supplied to the heat exchangers at a maximum of 180 psig. Each item of unit-level equipment has its own integral flow regulator, which will absorb any excess head. Pressure switches at each pump outlet actuate fault signals to the alarm panel and ORTS whenever either pump fails to provide the required head. Check valves in each pump outlet prevent reverse flow through a non-operating pump.

These design features are essential for isolating the Fire Control and Radar systems for maintenance or in the event of a catastrophic failure of the demineralizer/water cooler. Table 1 summarizes the design provisions for manual and automatic isolation.

Isolation in the event of a catastrophic failure is automatic, thereby eliminating the possibility of further equipment damage.

Summary of Significant Design Features

The design process described in this paper has resulted in a demineralizer/watercooler for AEGIS that is sized to meet performance requirements, has an adequate R&M safety margin, and includes the following features to reduce problems associated with existing shipboard cooling systems:

1. Both pumps are operated continuously although one pump can provide the required performance. Failure of one pump, therefore, will not cause an AEGIS failure. Also, a pump can be taken out of service for maintenance without interrupting AEGIS operation.

2. Only one heat exchanger is in use at a time. The second heat exchanger is available during maintenance or if a failure occurs.

3. Two temperature regulating valves are arranged in series so that a failure of either valve will not cause an AEGIS failure from out-of-tolerance water temperature. Failure of the thermal element in either temperature regulating valve shuts off 100% of the cooled water input to that valve. All of the temperature regulation is then accomplished in the other valve. The valves can be adjusted manually to "fine tune" temperature.

4. System faults are indicated visually and audibly at the alarm panel, and by signals to ORTS for rapid detection and isolation.

5. A hose connection and a seawater valve located immediately upstream from this connection provide an alternate method of supplying sea water to the heat exchanger in case of a casualty to the upstream sea water system. To accomplish this a jumper hose can be installed to an alternate sea water system, such as a fire main.

6. The duplex basket strainer contains two separate strainers, only one of which is required at any one time. This arrangement allows shifting and cleaning strainers without disturbing AEGIS operation.

7. On the expansion tank, both the gage glass and the low-level switch indicate low liquid level. The gage glass must be read by operating personnel. The low-level switch actuates the local alarm and sends a signal to ORTS. In addition, the low-level switch provides continuous dial indication of tank water level, as required by NAVSHIPS 0902-019-4000.

8. The separate supply and return lines for the MARK 99 Fire Control and AN/SPY-1 Radar systems are cross connected for redundancy in case of failure in any of the lines between the deckhouse and the electronic cooling equipment room. If a supply line is broken it is automatically shut off by the flow sensor. The cross-connection valve between the supply lines in the deckhouse can be opened manually and both systems will function normally, since they are supplied with cooling water from the remaining supply line. If a return line is broken it is isolated by manually closing a valve in the deckhouse and opening the cross-connected valve between the two return lines. This provides a common return for both systems. All supply and return lines are sized for the full flow of both systems.

9. Double tube-sheet heat exchangers have been specified for the AEGIS demineralized water system. This approach eliminates the contamination of demineralized water by sea water, due to leakage at the tube/tube-sheet connections, that occurs in units presently in use.

10. Flow-control orifices on the sea water supply lines are located a minimum of fifteen pipe diameters upstream from the AEGIS heat exchangers. This eliminates the highly turbulent flow in the heat exchanger bonnets that has caused local erosion.

11. The following corrosion resistant materials are used wherever possible for components and piping in contact with AEGIS demineralized water:

Copper	- MIL-T-24107, WW-T-775, WW-T-797, WW-T-799
Copper-Nickel	- MIL-C-15726E, MIL-T-16420
CRES	- Types 304, 316, 347
Bronze	- MIL-B-16540, MIL-B-16541

    Brass may be used only when components made of the above metals are not available. Brass will be limited to a nominal 15% zinc content, such as QQ-B-626, QQ-B-613.

12. To overcome the historical problems of haphazard installation of cooling systems with the attendant lack of access and incipient failure causes, the physical layout of the AEGIS cooling system has been predicted on maximum access to operating controls, hazard-free conditions of location, and optimum maintainability characteristics. The physical layout has been coordinated with the shipbuilders for the cooling areas of candidate ships for the AEGIS Weapons System including the DLG(N)-38 class ship. The layout will be controlled and maintained during construction by means of installation and interface control documents.

ACKNOWLEDGMENT

The final design of the AEGIS Demineralizer/Water Cooler was a result of the dedicated effort of NAVORD, APL and RCA personnel. Included are G. Buonagurio, APL's AEGIS program senior mechanical engineer; W. Powell, a designer of the AEGIS Cooling System who collaborated with the author in the incorporation of the many R&M design features; E. Britt, systems engineer and the author of "The AEGIS Cooling System," who was the source of the functional description and many of the design features and diagrams utilized in this paper; and G. Rogers, the cooling systems project engineer. These are but a few of the individuals who have contributed significantly to this effort.

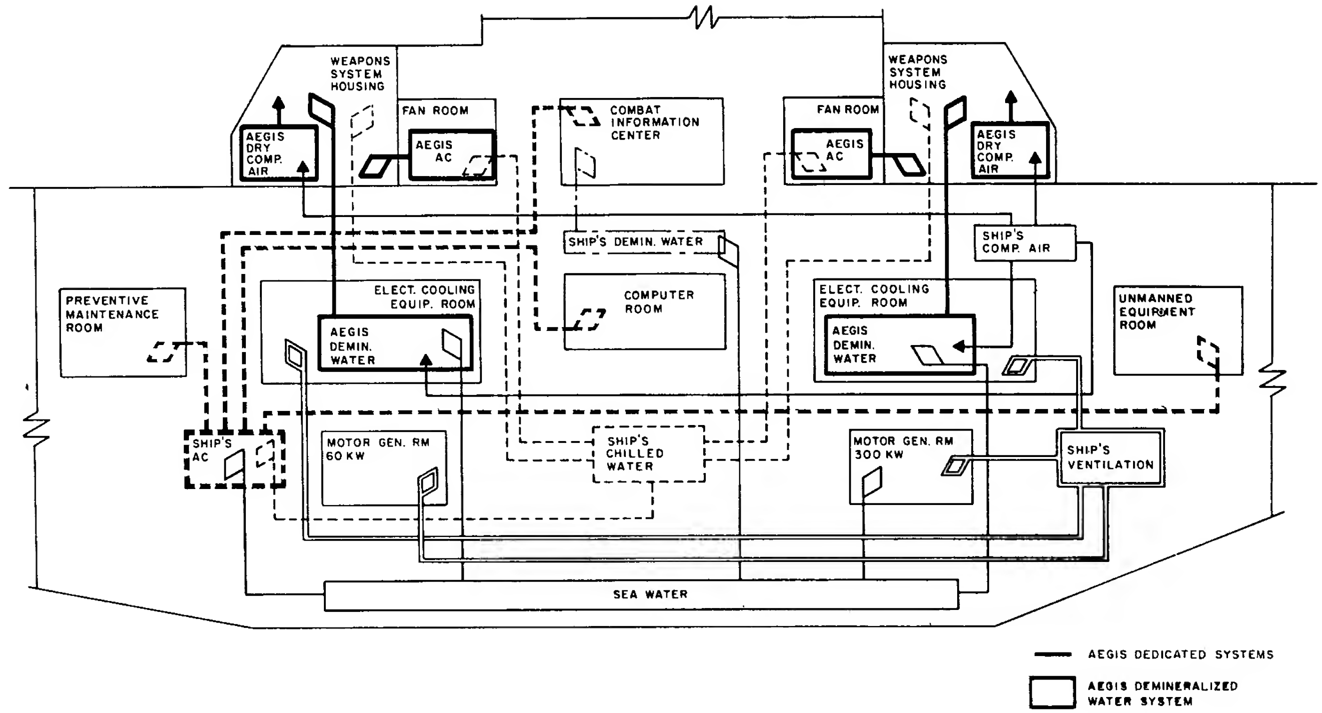


Figure 1. AEGIS Cooling System and Interrelated Ship's Services

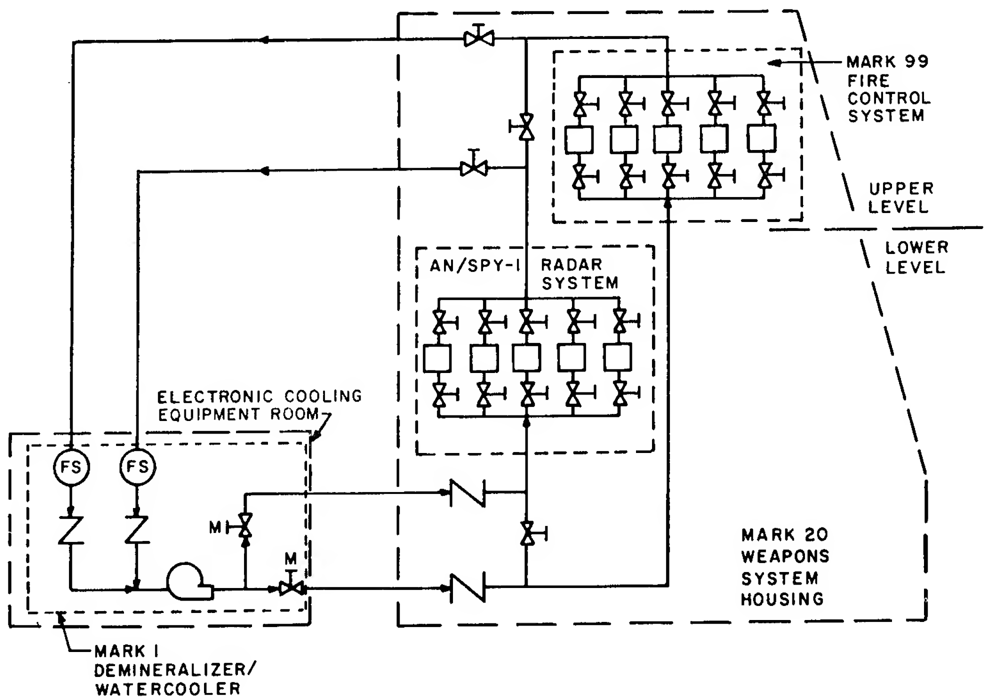


Figure 2. Demineralized Water System (Simplified Block Diagram)

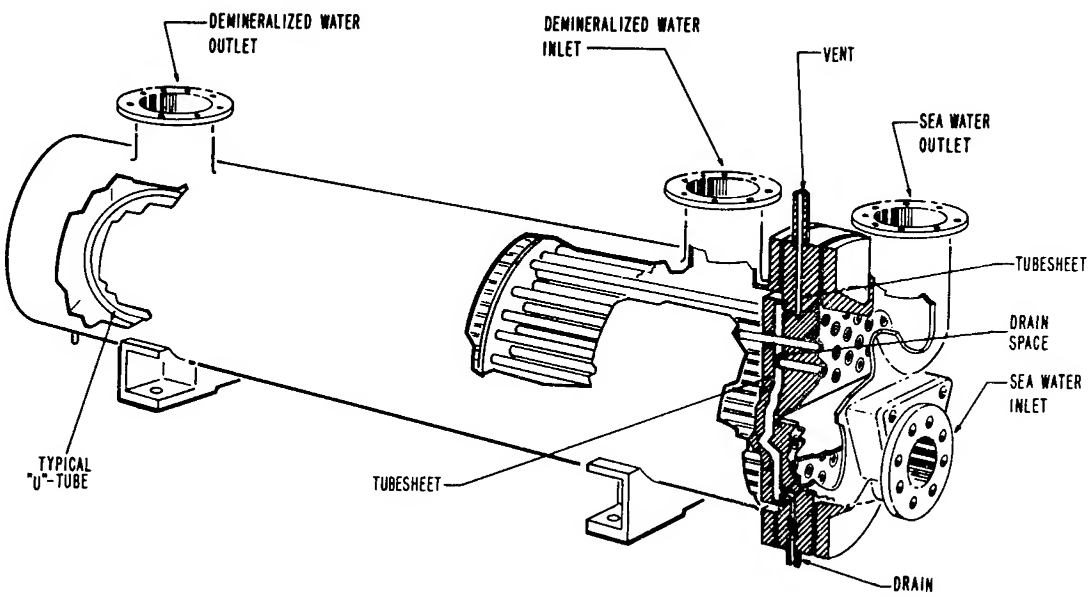


Figure 4. Double Tube-sheet Heat Exchanger

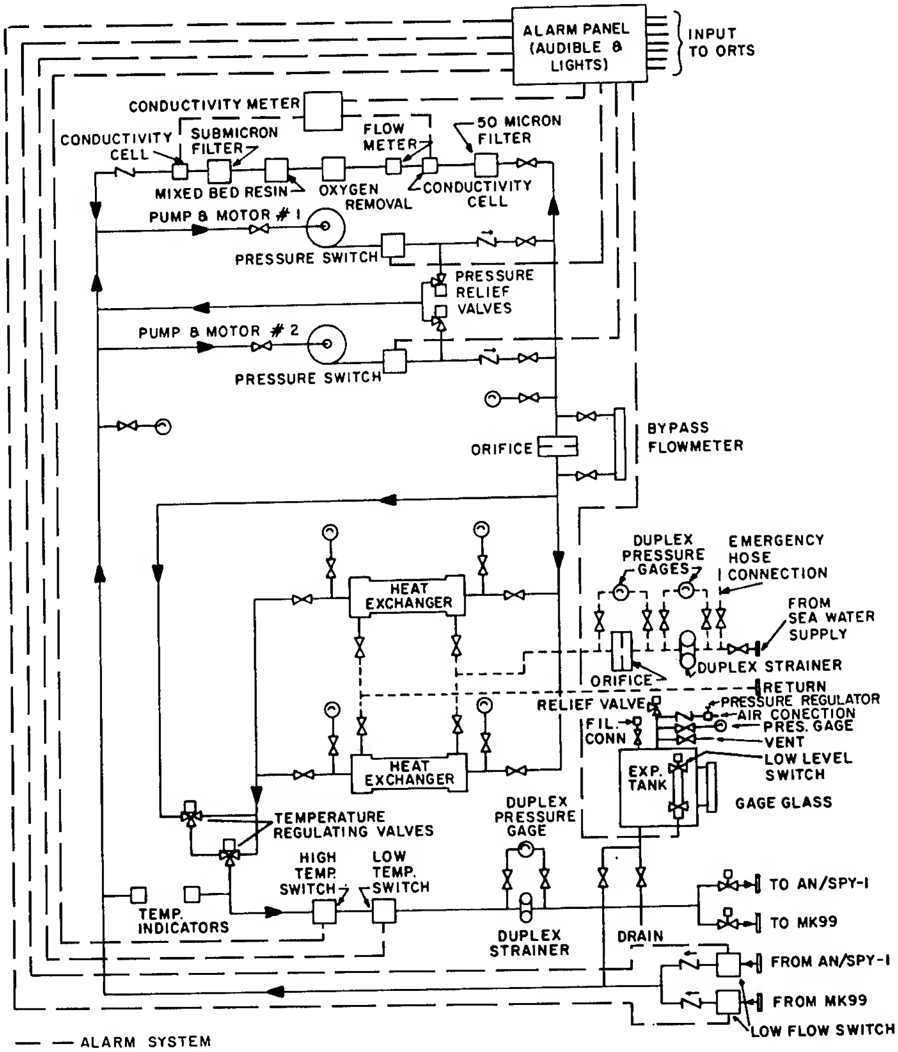


Figure 3. MARK 1 Demineralizer/Watercooler Showing Alarm System

TABLE 1. AEGIS WATER COOLING SYSTEM EQUIPMENT ISOLATION FEATURES

Activity Mode	AN/SPY-1 Isolation	MK-99 Isolation
Cabinet Maintenance Electrical Hydraulic	Manual Cabinet Shut-off	Manual Cabinet Shut-off
System Maintenance Electrical Hydraulic	Manual or Motorized Shut-off	Manual or Motorized Shut-off
Cabinet Catastrophic Hydraulic	Automatic Shut-off	Automatic Shut-off
System Catastrophic Hydraulic	Automatic Shut-off	Automatic Shut-off
Pipeline Catastrophic Hydraulic	Automatic Shut-off + Manual Cross Over	Automatic Shut-off + Manual Cross Over

AEGIS OPERATIONAL READINESS TEST SYSTEM - DESIGN FOR SYSTEM EFFECTIVENESS

Howard Boardman

William O'Leary
RCA, Missile and Surface Radar Division
Moorestown, N.J.

INDEX SERIAL NUMBER – 1048

SUMMARY

The AEGIS Weapons System is a Navy defensive missile system designed to shield the fleet against airborne threats of the late 1970's and 1980's. The achievement of high operational readiness with complex weapon systems, such as AEGIS, is dependent on the development of precise methods for measuring and controlling system status. In AEGIS, status control is supported by the Operational Readiness Test System (ORTS). The fundamental mission of ORTS is to extract data relative to the degree of operability of the system, process failure data for operation and maintenance purposes, and conduct those functions in such a way that the system availability and effectiveness requirements are met as specified. This paper shows how the ORTS is functionally related to system effectiveness and how availability modelling techniques have been utilized to provide a quantitative relation between system effectiveness and ORTS design parameters. The results of a few key parametric tradeoff studies are described to indicate the method used in developing ORTS requirements. Also included is a discussion of how the ORTS design enables these requirements to be achieved, and includes its functional interfaces, its major equipment elements, and its computer control functions.

INTRODUCTION

The achievement of high operational readiness with complex weapon systems, such as AEGIS, is dependent upon the development of precise methods for measuring and controlling system status. Failure to consider status control will result in operational readiness that is significantly lower than would be expected from inherent reliability and maintainability characteristics. In AEGIS, status control is supported by the Operational Readiness Test System (ORTS), which has been designed to approach this function from two directions:

- Maintenance of, or recovery to, higher operational capability after a malfunction or casualty
- Proper utilization of the weapon system and its equipments depending on their operational status.

The fundamental mission of ORTS is to extract data relative to the degree of operability of the system, process failure data for operation and maintenance purposes, and conduct these functions in such a way that the specified system availability and effectiveness requirements are met.

The uniqueness of ORTS lies not so much in its function, but in the engineering approach that integrates it into all elements of the AEGIS system. ORTS has its basis in functional requirements analyses, and its parameter values are established by RMA modelling techniques and computerized sensitivity studies. This paper will show:

- The relation between system effectiveness and the ORTS function
- The development of ORTS requirements
- The system implementation aspects of ORTS, including accomplishments.

ORTS DESCRIPTION

Figure 1 illustrates the basic functional interfaces established by ORTS with the AN/SPY-1 radar system. Highlighted are the key integral features, particularly those associated with the system computers. The basic functional loop consists of the radar system that is under the control of, and supplies data to, the same computer that processes ORTS data. The acquisition of test data is performed by computer-controlled addressing commands sent through the input/output (I/O) channels to the ORTS console, which in turn supplies control logic to acquire the data through a system of data/address busses routed through equipment cabinets (e.g. I/O buffer). Each cabinet contains a data acquisition assembly (DAA), which in turn can acquire data from approximately 512 test points. The net test point address capability for AEGIS is between 8000 and 10,000 data points.

The operational test data interface between the radar system and the radar system computer transmits mode and synchronization commands to the radar system, receives data from the radar system, and carries a two-way traffic of ORTS data. ORTS control programs in the radar system computer request scheduling of on-line simulation tests, and radar system output data is processed and returned to the ORTS functions within the radar system computer. All of these operations are under the control of the AEGIS Tactical Executive Program (ATEP), which orders schedules, establishes priorities and pre-emptions, and detects program fault data and interrupts.

AEGIS SYSTEM EFFECTIVENESS

The effectiveness of a system is a function of many factors, only one of which is the set of performance capabilities inherent in the design. Figure 2 is the AEGIS requirements structure, which shows the factors and interrelationships of factors that must be considered and controlled in order to achieve full system effectiveness.

The illustration shows that system effectiveness is not a single measure. For a multifunction and multimission system there may be several measures (i.e., SE_1, SE_2, $\cdots$ SE_N) that are important. Each system effectiveness measure would also have a unique set of functional relationships to the factors shown in Figure 2.

The first level of flow-down from system effectiveness covers:

- Performance parameters
- Operability parameters (i.e., availability and reliability)
- Command utilization.

Thus system effectiveness, as defined for the AEGIS Weapons System, can be expressed as

System Effectiveness = f (Performance, Operability, Utilization)

Another important factor is that the AEGIS reliability design approach is to take maximum advantage of equipment characteristics to minimize the impact of individual malfunctions on performance. This approach has as its goal zero downtime for the system, through the elimination of series links whose malfunction would abort AEGIS performance. This means that the design for, and evaluation of, system effectiveness must take into account all useful levels of performance with and without the existence of malfunctions. Command utilization refers to the use of readiness measurement and evaluation by all levels of AEGIS operational and shipboard command personnel to provide a basis for selection of operational procedures, configuration alternatives, and corrective program priorities.

SYSTEM EFFECTIVENESS AND ORTS

An effective ORTS system design requires:

- A clear definition of significant ORTS parameters
- Determination of the quantitative value of each performance parameter as necessary to support the system operability requirements
- The inclusion in the weapon system and system specifications of appropriate requirements relevant to ORTS performance.

Figure 2 gives the set of parameters in blocks that have been identified as those within the control of the ORTS function. These parameters may be placed into three categories:

- Fault detection
- Fault isolation
- System status reporting.

From Figure 2 it can be seen that these parameters have an impact on system effectiveness through their influence on availability and command utilization relationships.

DEVELOPMENT OF ORTS REQUIREMENTS

FAULT DETECTION

Fault detection refers to the sensing of a fault and the communication of the fault occurrence to the operation and maintenance personnel. Ideally, the detection of all faults is accurate and immediate. However, in practice, fault detection may depart significantly from this ideal: some faults may not be detected, test equipment failure may lead to false alarms, and detecting certain faults may be delayed. The design of an effective ORTS must consider all characteristics that have an effect on system availability. To this end, particular emphasis was focused on the development of an availability model that provides a quantitative relationship between availability and ORTS design parameters.

Inherent Availability Model

The model for the inherent availability of an equipment is normally defined as

$$A = \frac{MTBF}{MTBF + MTTR},$$

where:

A is availability

MTBF is the mean-time-between-failures

MTTR is the mean-time-to-repair.

This model is only valid for the following conditions:

1. All faults are immediately detected when they occur
2. Fault location and repair action is initiated as soon as the fault is detected
3. Failures are independent
4. No provision is made for the effects of monitoring imperfections, such as false fault indications.

The actual conditions encountered in practice may deviate considerably from the conditions for which this model is valid. Also, any deviations from these assumptions will generally lead to an availability that is significantly lower than that calculated using this model. Therefore, more general and/or more realistic models for availability, which are not dependent on the restrictions of these assumptions, must be developed. It will then be possible to obtain a better understanding of the impact of ORTS characteristics on availability. Realistic requirements can then be applied to ORTS equipment, as well as the usual MTBF and MTTR requirements, to achieve a specific level of availability.

Realistic Availability Models

To provide a basis for modelling availability, faults may be classified and grouped into categories depending on when they are detected. Those faults that would normally be detected within 1 hour of occurrence are classified as monitoring-detected faults. Those faults that would be detected periodically (i.e., less frequently than once an hour) by test are classified as operability-test-detected faults. All faults not covered by monitoring or operability tests are classified as undetected faults. The detection of faults is essential to status control, which includes the initiation of corrective maintenance action and system utilization.

It is quite possible for a fault to be detected by both monitoring and operability tests. In these cases, the fault is classified as being among those detected by monitoring since monitoring should almost always detect the fault before the operability test. By definition, the three categories of faults are mutually exclusive. The total failure rate of an equipment can be expressed as the sum of the failure rates associated with each of the three fault categories:

$$\lambda_T = \lambda_m + \lambda_o + \lambda_u ,$$

where:

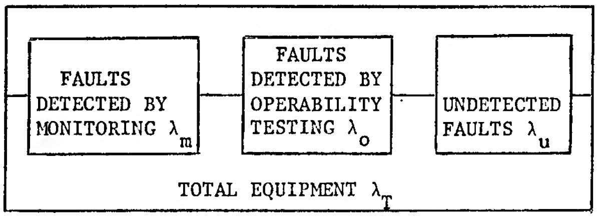
λ_T is the total failure rate for all operational equipment faults measured in faults per hour

λ_m is the failure rate associated with those faults detected by monitoring

λ_o is the failure rate associated with those faults sensed by operability testing

λ_u is the failure rate for undetected faults.

The equivalent reliability block diagram for an equipment may then be drawn as



The approach taken is to model the availability for each fault category and then to define the total availability as the product of each of the availabilities. Thus,

$$A_T = A_M \, A_O \, A_U ,$$

where:

A_T is the total equipment availability

A_M is the availability for those equipment elements with faults detected by monitoring tests

A_O is the availability for those equipment elements with faults detected by operability tests

A_U is the availability for those equipment elements with faults undetected by monitoring or operability tests.

Monitoring Model - The derivation of the monitoring model for availability (A_M) rests on the definition of various combinations of conditions of the operational and monitoring equipments. The operational equipment can be good or bad and the monitoring equipment can be good or bad. In addition, the bad conditions of the monitoring equipment can be further subdivided depending on the fault effects in monitoring indications. As an example, it is possible that a monitoring equipment fault leads to a false indication of operational equipment failure without a self-test indication that the monitoring equipment has failed. This fault effect has the effect of initiating a false repair on the operational equipment.

The monitoring model is not presented here because of its complexity. However, it is functionally defined as:

$$A_M = f\,(\lambda_M, \lambda_i, MTTR_M, MTTR_i) ,$$

where:

λ_M is the failure rate of the monitoring-detected faults in the operational equipment

$MTTR_M$ is the meantime to repair these operational equipment faults after detection

λ_i is the failure rate of monitoring equipment faults depending on failure effect (e.g., λ_2 in the model is the failure rate for those monitoring equipment faults that are detected as monitoring equipment faults by self test)

$MTTR_i$ is the mean time to repair a monitoring equipment fault with the "i_{th}" failure effect

For ORTS design control, the monitoring-availability model variables have been defined in the following ORTS parameters:

Fault detection coverage

Test accuracy

Monitoring equipment self-test coverage

Monitoring equipment reliability and maintainability

Monitoring equipment fail-safe characteristic.

Operability-Test Model - In the monitoring model, the fault detection time is assumed to be approximately zero due to the essentially continuous nature of monitoring. However, in the operability-test model, a fault could exist for a considerable period prior to the next operability test. These effects have been included in the derivation of the operability-test model for availability. Parameters used in the operability-test model are defined as:

λ_o is the failure rate of operational equipment faults detectable by operability test

T is the test interval or the time between the completion of one operability test and the initiation of the next test

T_o is the length of operability test period downtime. Note that T_O may be less than the total time required for test if a portion or all of the test is accomplished without incurring downtime

$MTTR_o$ is the mean time to repair operational equipment faults detected by operability test

E is the operability test efficiency, or the probability that an equipment that is detectably bad will be called bad by a single application of the test

α is the false alarm probability, or the probability that an equipment that is detectably good will be called bad by a single application of the test.

These parameters are then related to availability through

$$A_o = \frac{\left[P_u \quad \frac{1}{\lambda_o}\left(1 - e^{-\lambda_o T}\right)\right]}{\frac{1}{2} P_u \, MTTR_o \, e^{-\lambda_o T}(\alpha E - E)\left(1 + e^{-\lambda_o T_o}\right) + T + T_o + MTTR_o(E + \alpha - \alpha E)} ,$$

where:

$$P_u = \frac{\frac{1}{2}E\left(1 + e^{-\lambda_o T_o}\right)}{1 - (1 - E)e^{-\lambda_o (T+T_o)}} .$$

This model provides average availability for faults detected by operability test. A typical representation of expected availability through time would be that shown in Figure 3. Three cycles are shown to illustrate the meaning of the parameters, rather than actual performance.

The first cycle is shown starting with unity availability at zero time. The exponential drop in vailability over each of the test intervals, T, is a

reliability-type function associated with the failure rate, λ_o, of operational equipment faults detected by operability test. The test period, T_o, which follows each operability test interval, T, is shown as zero availability.

The second cycle shows a repair time, MTTR, following the test period. The availability level, P_u, at the beginning of the third cycle is the expected long-term probability of being up at the beginning of a cycle.

For ORTS design control, the operability-test availability model variables have been defined in the following ORTS parameters:

Fault detection coverage

Test periodicity

Test efficiency

False alarm probability

Time to recover from test

Test duration.

Undetected-Faults Model - Undetected faults are defined as any degradations in equipment performance below satisfactory operational limits that are not detectable by the monitoring or operability test. It is assumed for the unsensed-faults availability model that more extensive tests can be conducted at the shipyard to detect the existence of undetected faults, and that the equipment can be returned to full performance status. Further, the assumption is made that undetected failures have a constant failure rate, λ_u.

The availability model for undetected faults is given by:

$$A_u = e^{-\lambda_u t}$$

where:

λ_u is the failure rate for undetected faults

t is the operating time since the last complete shipyard test

Figure 4 illustrates the characteristic of availability as a function of time for the undetected-faults model.

Parametric Tradeoffs - The establishment of allocated requirements for the ORTS parameters was based on the results of parametric trade studies with the availability models coupled with experience gained on other systems. The results of these trade studies led to a feasible set of ORTS parameter values that were incorporated into the system specifications. It is of interest to review a few of the key parametric tradeoffs and the conclusions that were drawn as a result.

Fault Detection Coverage - The most critical ORTS parameter from an availability standpoint is fault detection coverage. As an example of fault detection coverage impact, an equipment was selected with an MTBF of 200 hours and an MTTR of 2 hours. This results in a basic inherent availability of 0.99. Figure 5 shows the sensitivity of availability, A_u, to undetected faults as a function of time. The figure includes curves for 0.1%, 1.0%, and 5.0% undetected faults. The beginning of the period may be interpreted as being just after completion of a comprehensive in-port test and system restoration from all faults. The availability is a function of the probability of a fault occuring after this test and remaining undetected by monitoring or operability tests. This represents a direct degradation of inherent availability. It is apparent from the results in Figure 5 that the design for ORTS must drive to a 100% coverage of all faults. Even low percentages of undetected faults would significantly degrade the 0.99 inherent availability in less than 100 hours of system operation. All AEGIS system specifications contain the requirement that 100% of all faults be detectable by the combination of monitoring and operability tests.

Monitoring Equipment Self Test - In the development of the availability models it was recognized that monitoring equipment failures could lead to various false indications:

- The monitoring indicates good while a fault exists in the operational equipment.
- The monitoring falsely indicates a failure in the operational equipment to start an unneeded repair action.

These false indications not only have an adverse affect on availability but they tend to lower maintenance operator confidence in the ORTS system. Self test refers to the built-in capability of the monitoring equipment to distinguish between operational-equipment faults and monitoring-equipment faults.

Figure 6 shows the two extremes of 100% self test and no self test as a function of the ratio of monitoring equipment MTBF to operational equipment MTBF. With a 100% self test we are always able to know whether a fault indication has been initiated by a fault in the monitoring equipment or the operational equipment, so that the reliability (MTBF) of monitoring equipment has little impact on availability. Conversely, for no self-test, equipment availability is extremely sensitive to the MTBF of the monitoring equipment. The no-self-test curves in Figure 6 are only for the "false alarm" case; that is, when a fault in the monitoring equipment leads to a false repair of operational gear. The companion case of the monitoring equipment failing to indicate an operational equipment fault results in a similar set of curves.

It was concluded that availability degradation is minimized by high self-test coverage and a high MTBF for the monitoring equipment. To reflect these results system specifications contain a requirement for a 90% self-test coverage and that the monitoring equipment be at least 10 times the reliability of the operational equipment. Taken together these requirements assure a minimum impact on availability.

Operability-Test False-Alarm Probability - Figure 7 shows the relationship of availability to operability-test false-alarm probability as a function of the time interval between operability tests. These curves indicate that it is very important to hold the false-alarm probability to a minimum and that even with a zero false-alarm probability an extended period between operability tests would degrade availability. Consequently the false alarm probability, α, must be very close to zero and the amount of time between operability tests must be small. To reflect these results, system specifications contain requirements that the false-alarm probability be held to less than 0.005 and that the operability test interval be held to 8 hours or less.

FAULT ISOLATION

Fault isolation is the only ORTS characteristic that can be reflected, through its impact on MTTR, in the inherent availability mode. Fault isolation parameters of direct interest to the ORTS designer are:

Fault isolation time

Number of replacement items at the level of isolation

Maintenance skill levels required.

Previous Navy experience indicated that it was essential for all AEGIS maintenance be within the skill capabilities of the average Navy maintenance personnel. Therefore, ORTS has been designed to provide automatic or semiautomatic fault isolation to the lowest replaceable unit with unambiguous indication to maintenance personnel of action to be taken. In essence, the bulk of the logic of fault isolation is designed into ORTS rather than placing dependence for this complex task on the maintenance man.

By this approach fault isolation times are held to a minimum. Those equipment areas covered by automatic fault isolation take less than a minute, and areas covered by semiautomatic fault isolation take only a few minutes. With ORTS, the contribution of fault isolation time to MTTR is relatively small.

Finally, the number of items in the line replaceable unit (LRU) can have an important bearing on remove-and-replace times and aboard-ship sparing requirements. All system specifications include the objective that the average number of items in an LRU be 5 or less. In the case of a digital equipment cabinet this would mean that the fault is isolated to 5 module cards. The subsequent repair action is then directed to that group of 5 module cards.

SYSTEM STATUS REPORTING

System status reporting is a vital function, of command utilization, see Figure 2. Command utilization is of particular importance to systems, such as AEGIS, that have been designed to be resilient to faults by providing multiple levels of useful performance. If malfunctions exist within the system,the status reporting function of ORTS becomes vitally important to:

- Initiate corrective maintenance to restore performance to a higher level
- Provide information necessary to initiate system reconfiguration
- Provide information to make effective use of the remaining performance capability.

The AEGIS design approach capitalizes on this ORTS function through a centralized status reporting for all AEGIS systems and equipments.

If multiple malfunctions exist at the same time, centralized status reporting affords the opportunity to assign priority in the dispatch of maintenance personnel to the trouble spots. If a computer should fail, the status reporting opens the possibility for automatic or manual reconfiguration of the remaining computers to continue weapon system operation at some reduction in capability. If a major weapon system element, such as an illuminator, should fail the status reporting is used directly to bypass the failed unit, and use only the operating illuminators until full performance is restored. If a functional capability, such as the moving target indicator (MTI), is lost in the AN/SPY-1 radar signal processor,the radar would continue to detect and track targets making use of other designed-in capabilities.

AEGIS SYSTEM IMPLEMENTATION ASPECTS OF ORTS

The modelling techniques previously outlined for relating ORTS performance to AEGIS availability, identified certain "key parameters" that establish the performance cornerstone of ORTS. Certain of these parameters, such as fault detection coverage, are achieved by a combination of ORTS and AEGIS operational equipment. Other parameters, such as ORTS self test, monitoring accuracy, test efficiency, false-alarm rate, and test periodicity (scheduling), are directly controllable by the ORTS hardware-computer program design. While these latter parameters are allocated in a controlled fashion by a flowdown of specifications (AEGIS System, ORTS Segment, ORTS hardware, and computer programs), it is interesting to note that in many cases the point at which the implementation of one of these parameters is most visible may well be in what would appear as a minor portion of the design layout (e.g., a circuit board, or an inconspicuous portion of the coding of an ORTS program module). In other parameter implementations, the key implementation level may be in the overall functional requirements of a complete computer program module, in the functional interface between ORTS programs and the common executive program of the resident computer, or in the combined performance of the elements in the ORTS test point data acquisition system. In the ORTS system engineering process, we call this "system requirements sensitivity at all levels of design," and several of these key points of sensitivity are briefly analyzed in the material that follows.

THE INTEGRAL NATURE OF ORTS

While system integrity is a necessary part of a system design approach, ideally it has only a qualitative effect on ORTS performance achievement. Supposing that equipment space was unconstrained, power and weight were not vital, computer core storage was unlimited, design could overcome any interface problem, and intercomputer I/O message traffic was unlimited, possibly enhanced by memory sharing techniques, then the integral nature of ORTS within the AEGIS System would not be a dominant factor. However, in practice, all of the above considerations, and many others, have guided system design studies and affected implementation tradeoffs in the direction of an integral relationship between ORTS and the equipments with which it operates. The nature of this system integrity allows us to highlight some of the more significant configurations and capabilities that have evolved, and to understand their contribution to ORTS/AEGIS performance.

Functional Implementation Highlights

Figure 8 illustrates a typical functional configuration including the major elements of ORTS. The ORTS test and monitor (T&M) console is the control and display position manned by a maintenance supervisor as shown in Figure 9. The T&M console provides five-state rear projection status display panels, an alphanumeric CRT and keyboard interfacing with the computer, and a high-speed printer to output hard copy on any data appearing on the CRT. The T&M console also contains address control electronics for routing interrogation commands to, and accepting data from, the test acquisition modules (TAM) sensor cards mounted within DAA assemblies located in each equipment cabinet.

The ORTS-developed DAA, shown in Figure 10 is a unit approximately the size of a large loaf of bread, containing addressable TAM cards connected to an external data/address bus. DAA's are mounted in system equipment cabinets, and connected to one common data/address bus running from the T&M console. Each TAM provides address control logic to handle up to 16 test points, and also converts the measured values to a serial digital format and transmits it back on the data/address bus. Each test point has a unique address, and the DAA provides the multiplexed receiver/transmit functions to allow the computer to interrogate any test point individually and in an established sequence. The DAA TAMS are capable of interfacing with analog, serial digital, or parallel digital data from equipment test points.

The AN/UYK-7 computer is the principle control element of ORTS. Figure 8 indicates five of its critical control functions. The test scheduling (with periodicity established according to equipment failure sensitivity guidelines provided by the system designers) is controlled by the computer and fulfills the system requirements for overall periodicity of monitoring and operability testing. The spectrum of test schedules runs from critical tactical elements tested every 20 seconds, to those assigned a 4-to-8-hour periodicity. This is a somewhat radical departure from the test design of existing systems and is achieveable by virtue of the "on-line" nature or interleaving of test programs and tactical operations (see Figure 8). Within some tactical constraints, this provides ORTS with an unlimited choice of periodicities, within which essentially all test requirements can be managed.

Another subtlety of the integral nature of ORTS within system computers is the mutual access and exchange of data. Commonly fed buffers and tables accumulate equipment file and functional status information within the computer, and provide ORTS with a powerful data base for operability analysis.

Contributing to ORTS ability to meet false-alarm requirements is a test criteria that calls for M fault indications out of N tests (M/N) before a failure report is shown at the T&M console. This technique to limit false-alarms is applied to most monitoring and simulation tests. The sensitivity curves of availability vs false bad report, shown earlier in Figure 6, indicate the impact of false alarms on availability.

Data accuracy and self-test requirements are the two remaining dominant control functions, (see Figure 8).

Within the classification of self-test there are:

1. End-around tests of the computer-to-ORTS console I/O interface, which are run periodically.

2. Calibration tests of the "TAM" sensor cards within the DAAs. Two types of tests are run here. For TAMs that accept analog (dc) data, a self-contained calibration voltage is applied to its input and the computer checks the output. For TAMS that accept digital data, a fixed "I", an "O" pattern of n bits is inputted and the serial transmission to the computer is checked; in effect this process checks out the complete data acquisition system through the console.

3. The accessing of DAA data, via the console address control electronics, to the computer is subjected to parity checks. In addition, there are no two TAM addresses closer numerically than two, which minimizes the possibility of false addressing. (This latter capability represents a system compromise, since allowing the numerical spacing of two essentially cuts in half the quantity of real TAM addresses useable. Systems without the stringent requirements of the AEGIS shipboard environment can increase their data accessing capability by reducing this requirement.)

The self-test functions of ORTS are closely related to the data accuracy requirements. The calibration self-test procedures on TAM data ensures that small offset variances in TAM outputs are compensated by the software and do not show up as a bias that could affect a thresholding decision in the program. Incorporating this capability allows the total data acquisition system to operate under a 99% data accuracy requirement.

Supporting overall system equipment configuration management in casualty modes, the ORTS console has reassignable computer I/O channels, so that with appropriate reloading of ORTS program modules an active computer can pick up the ORTS functions of a machine that is down for repair.

CONCLUSIONS

The achievement of high operational readiness with complex weapon systems, such as AEGIS, is dependent on the development of precise methods for measuring and controlling system status. In AEGIS, status control is supported by the Operational Readiness Test System (ORTS). The engineering approach to the ORTS design was that of integration into all elements of the AEGIS system. Quantitative requirements for ORTS are established to optimize ORTS performance as it directly impacts system effectiveness through its impact on system availability and command utilization. The ORTS and weapons systems equipments have been designed to achieve the ORTS requirements for high AEGIS operational readiness and are being readied for the test program to provide proof of accomplishment.

ACKNOWLEDGEMENT

The authors wish to acknowledge the many contributors to the design of the system and in particular E. Wares, the Navy ORTS project engineer, and R. Howery, the RCA ORTS project manager.

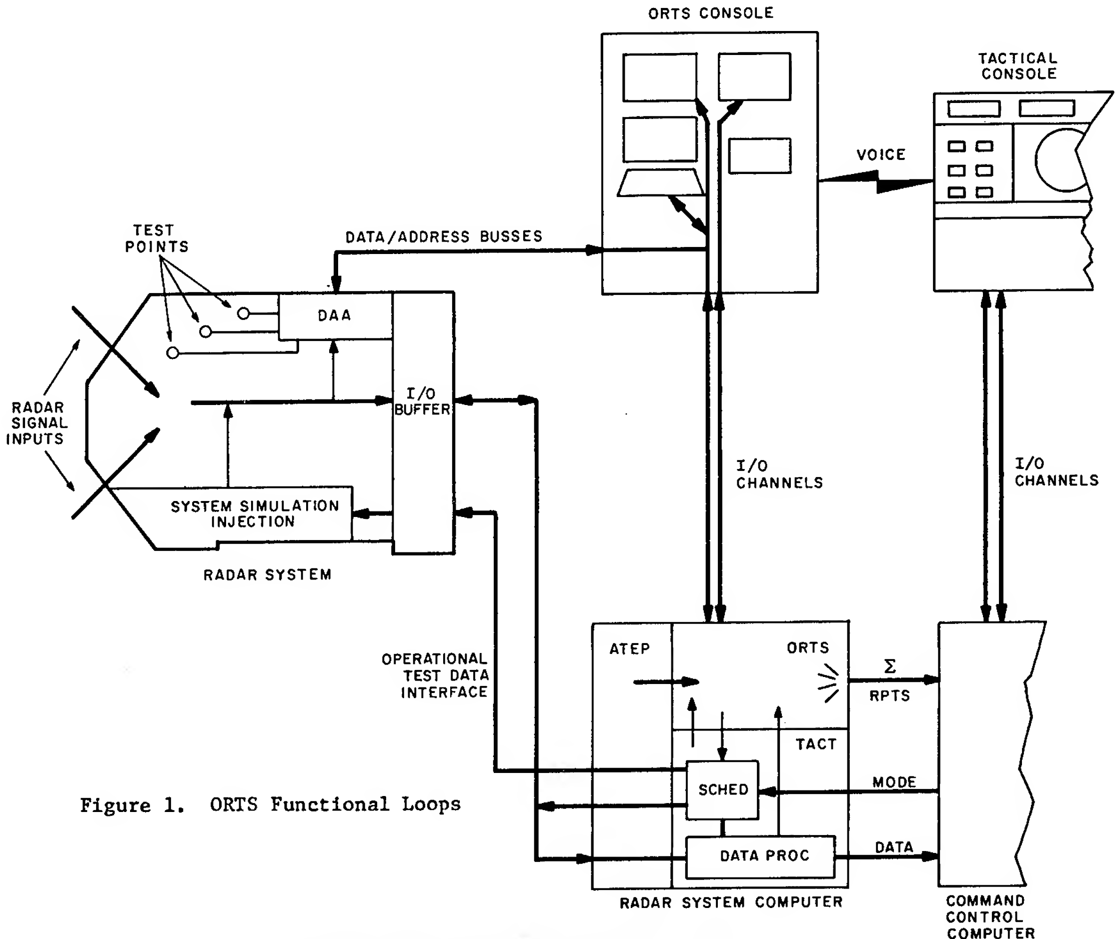


Figure 1. ORTS Functional Loops

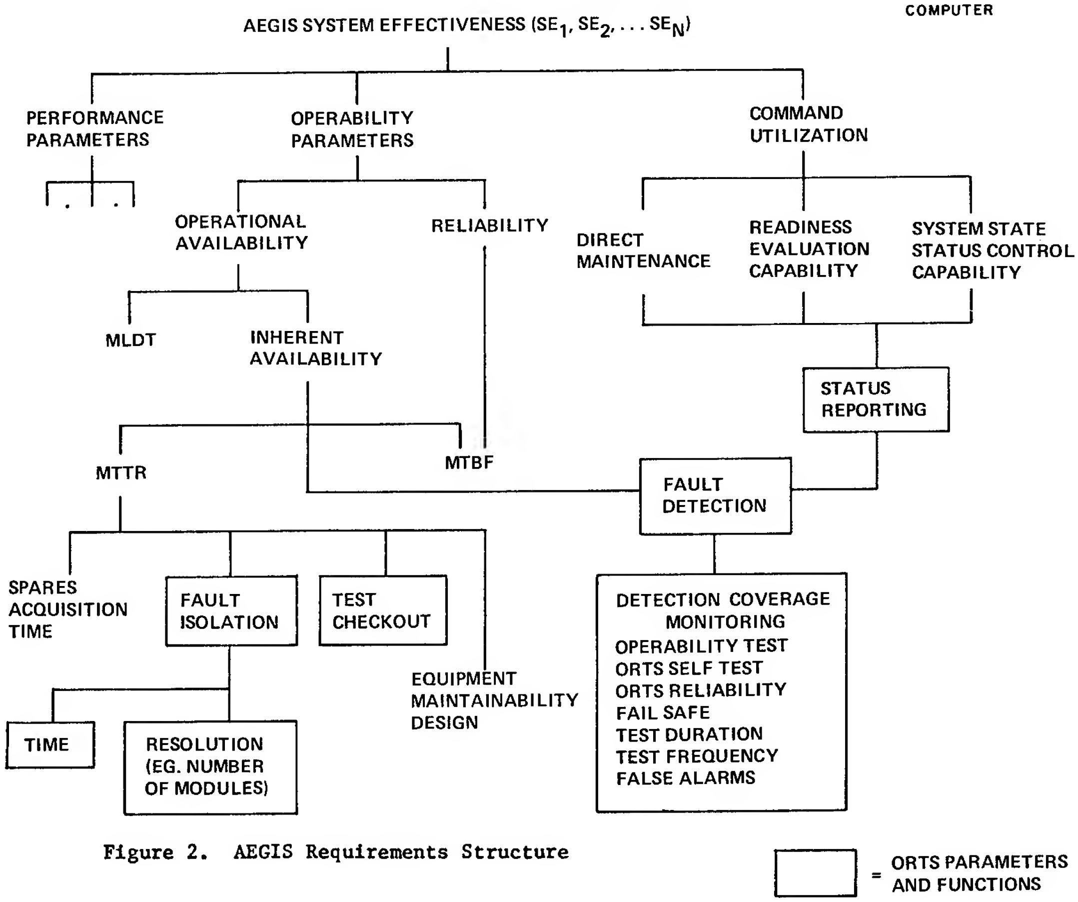


Figure 2. AEGIS Requirements Structure

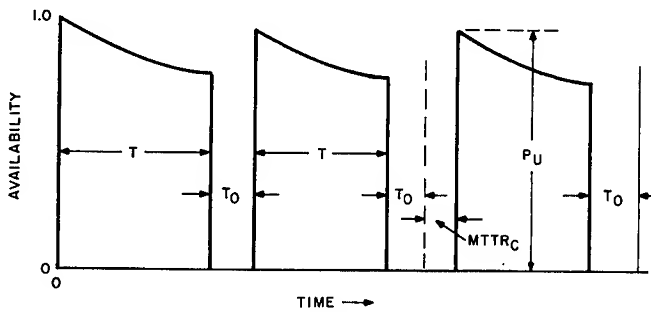


Figure 3. Typical Cycles for Availability as a Function of Time for Operability-Test-Detected Faults

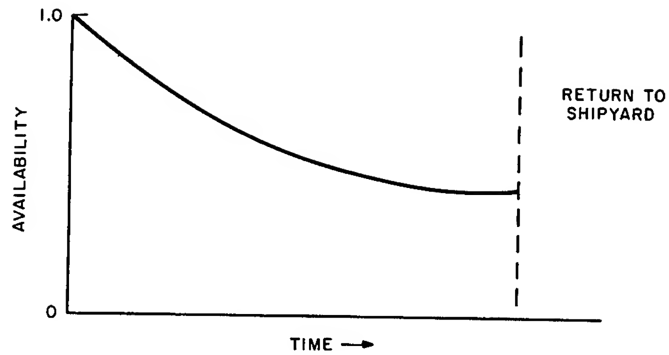


Figure 4. Operational Readiness as a Function of Time for Undetected Faults

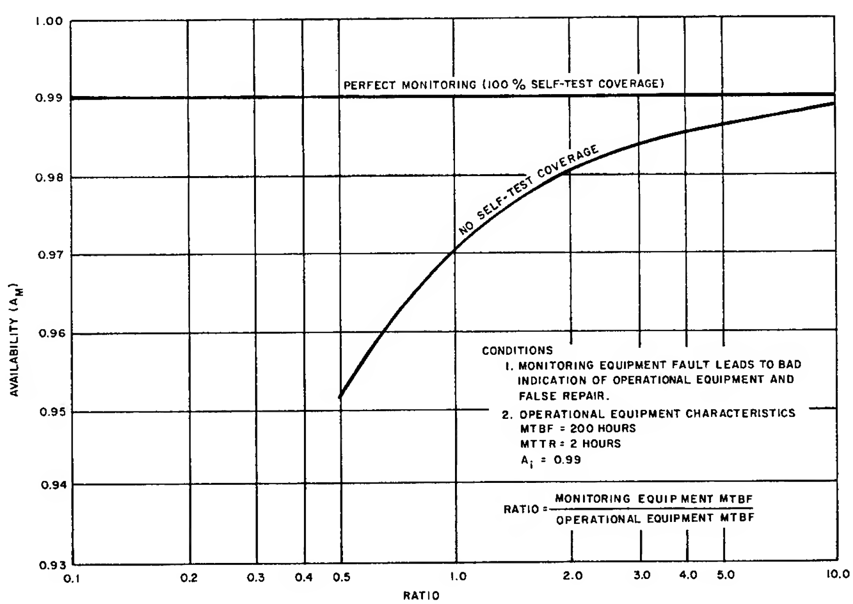


Figure 6. Sensitivity to Monitoring Equipment MTBF and Self-Test Coverage

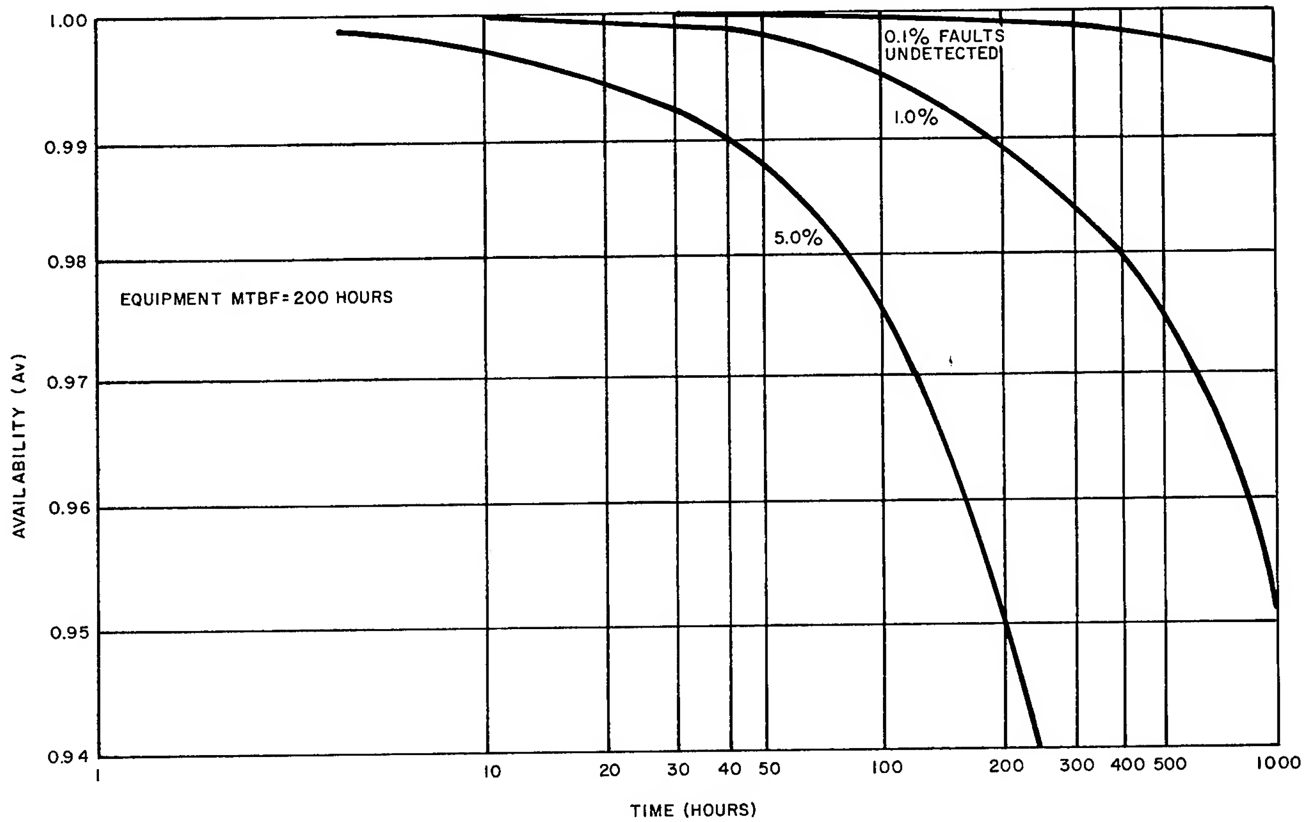


Figure 5. Availability Sensitivity to Undetectable Faults

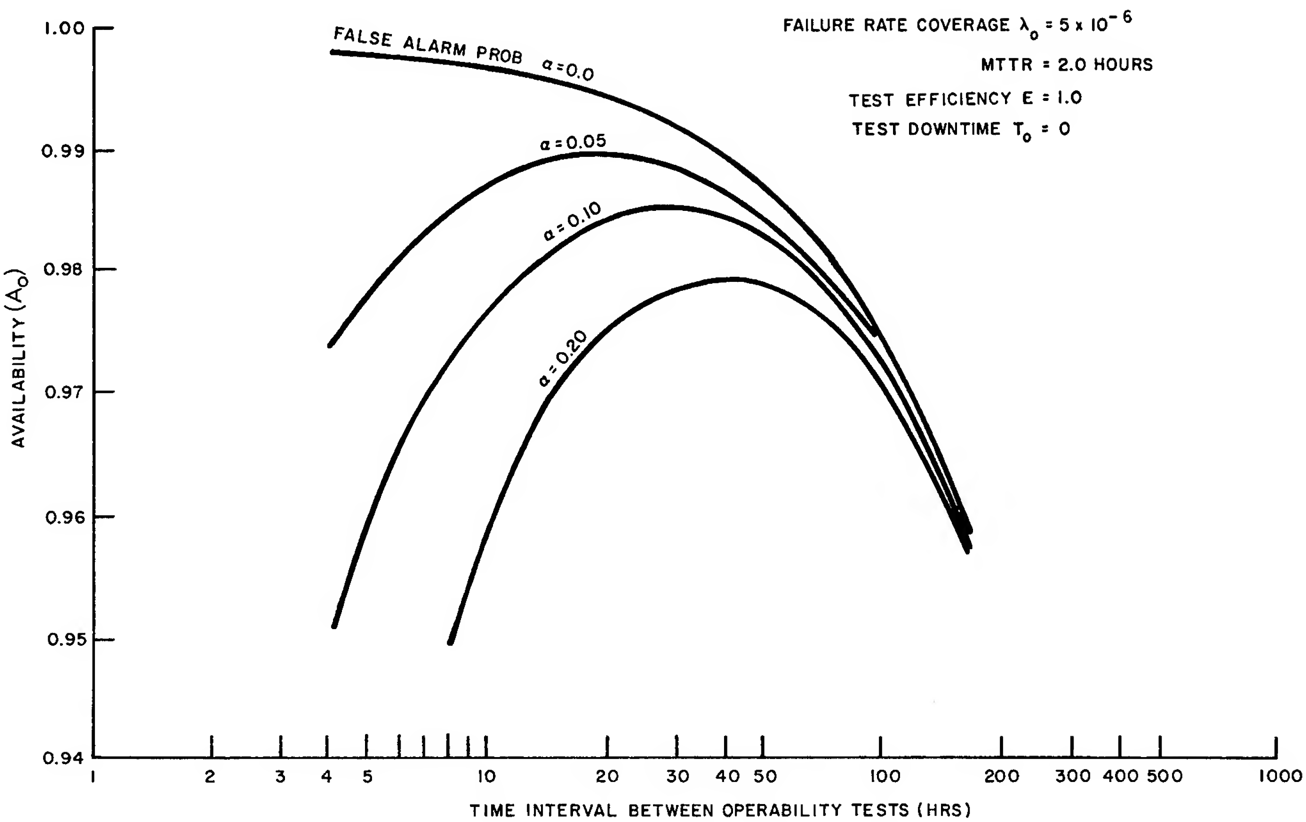


Figure 7. Sensitivity to Time Interval Between Operability Tests as a Function of False-Alarm Probability Curves

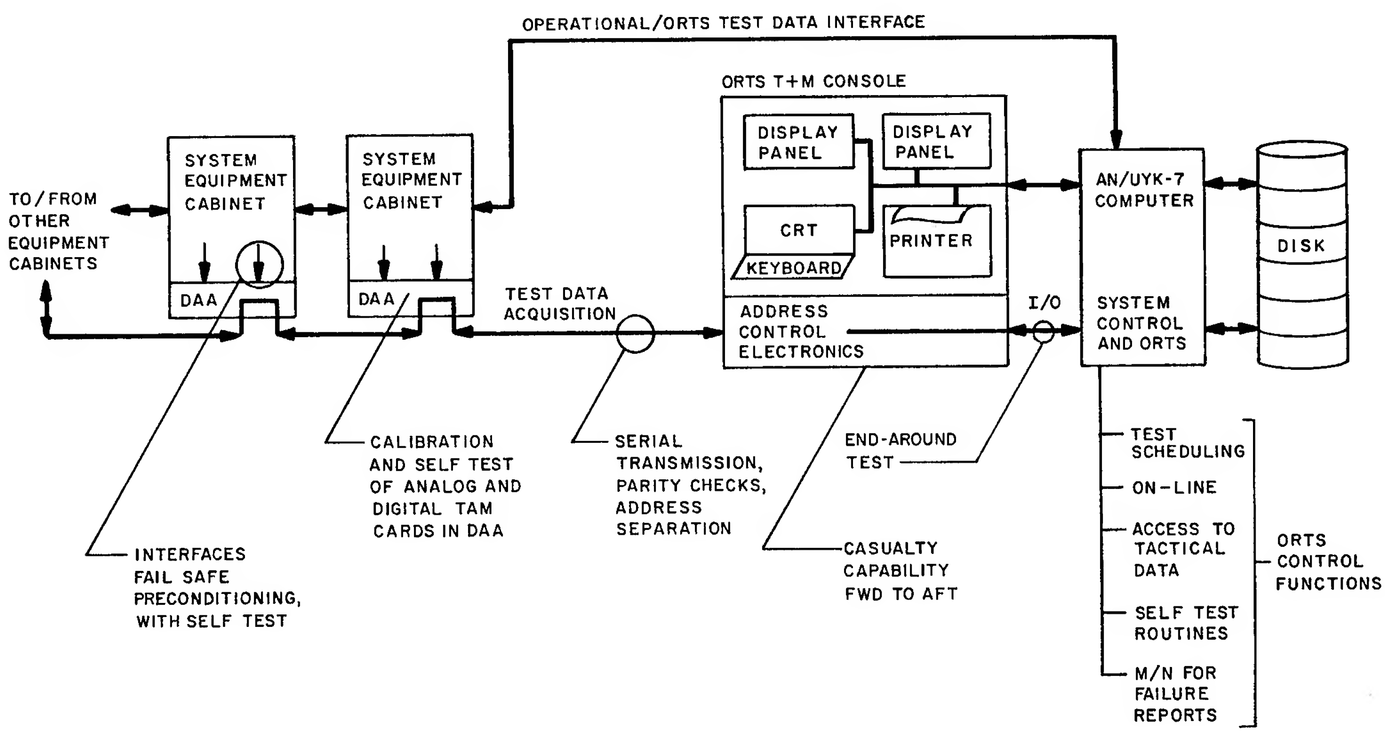


Figure 8. Basic Elements of ORTS Functional Loops

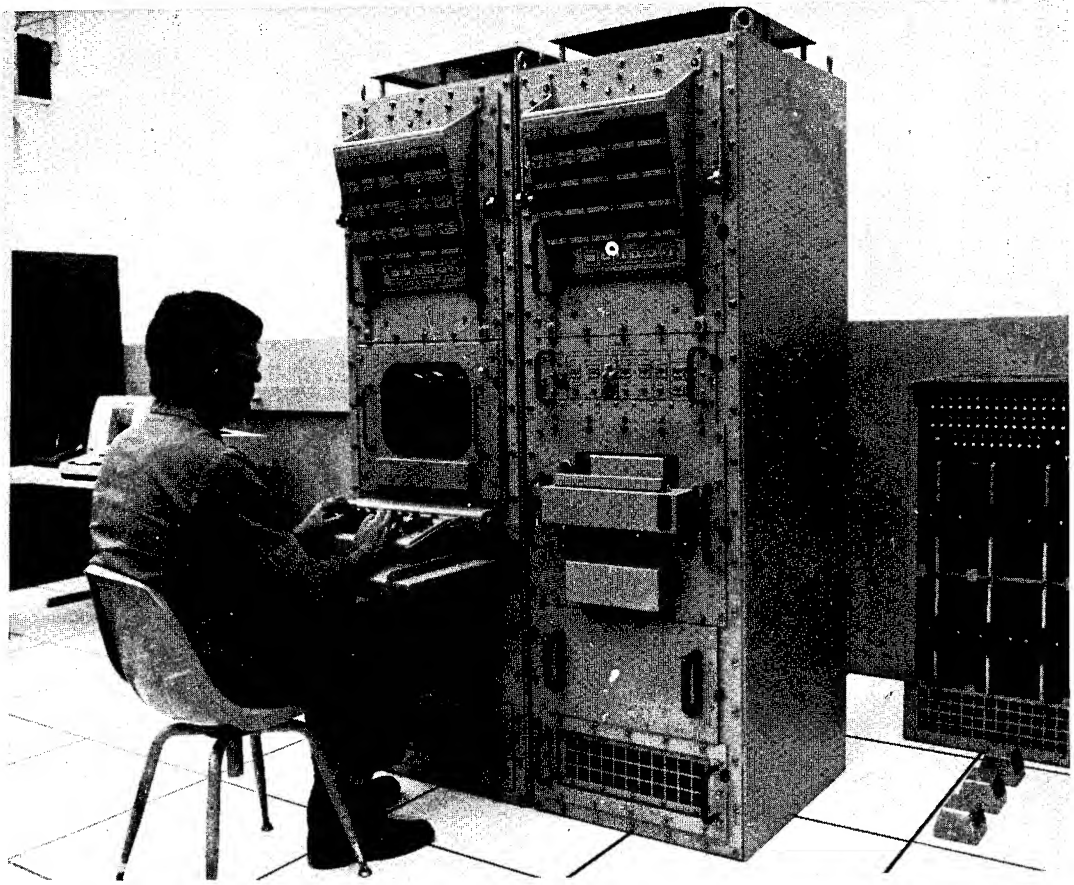


Figure 9. ORTS T&M Console in Operation

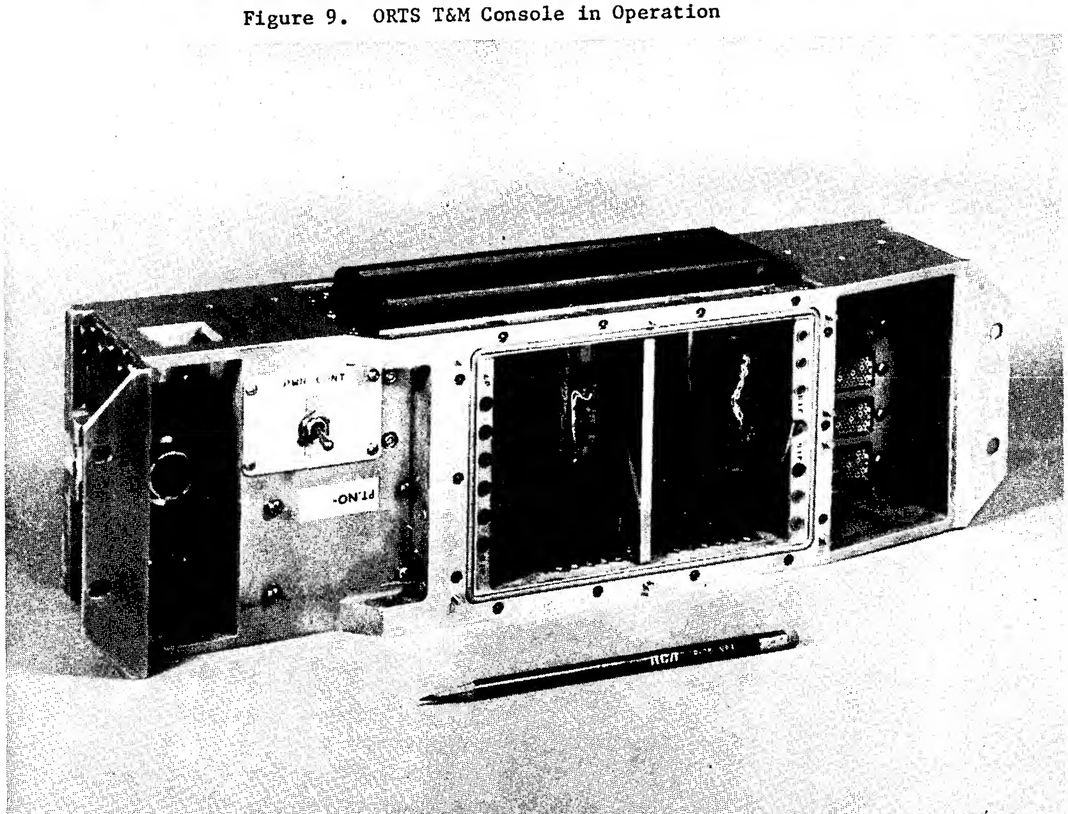


Figure 10. An ORTS Data Acquisition Assembly - Cover Open Showing Receptacles for TAM Sensor Cards

APPROACH TO RELIABILITY FOR THE SM-2 MISSILE

J. C. Bear
General Dynamics
Pomona, California

The SM-2 Missile

The SM-2 Missile, now undergoing engineering development, is the second version of the Standard Missile, and is the successor to the TERRIER and TARTAR missiles that have served as the primary air defense weapon for the U.S. fleet for the past 20 years (Figure 1).

SM-2 is being designed for use with the AEGIS Weapons System and also to upgrade the capability of the existing TERRIER and TARTAR ships in the fleet today (Figure 2). The missile will have two basic configurations: (1) the medium range (MR) round, which will employ a dual-thrust rocket motor and will work with the AEGIS and TARTAR weapons systems; and (2) the extended range (ER) round, which will incorporate a booster and a sustainer rocket motor and will be compatible with the TERRIER Weapons System.

The SM-2 missile will provide significant increases in performance, such as engagement range, kill probability, fire power, and ECM immunity. Although the AEGIS, TERRIER, and TARTAR fire control systems differ, the SM-2 missile will provide these improvements to all three types of ships by means of a common guidance and autopilot system that is functionally interchangeable between the MR and ER configurations.

The Challenge

Achieving these performance improvements presents a major challenge to the SM-2 program with respect to reliability. The total number of electronic parts will increase by about 70 percent (Figure 3), integrated circuits will be increased by a factor of seven, and large scale integration will be introduced. Because of this increased complexity, the flight reliability of the SM-2 must be greater than that of its predecessor, SM-1. Also, cost will be treated as a design parameter of very high priority.

This is not a new challenge. Missile evolution has been relentlessly prodded by the increasing sophistication of the threats. New threats require the addition of new performance features. Incorporating these new features requires space, and the space is made available by the continuing trend toward miniaturization. Therefore, the number of parts continually rises, and with the increase in complexity comes an increase in the source of variability and potential failure and unreliability.

Fortunately, miniaturization brings with it some reliability benefits; i.e., the newer, smaller parts are more reliable per function than their predecessors. However, miniaturization does not improve reliability quite as fast as it creates space (for more parts); therefore, the missile engineer is hard pressed just to keep even. With SM-2 the goal is not just to stay even by maintaining the same reliability as SM-1, but to continue the steady reliability growth that has characterized this family of missiles in the past.

The Approach Taken

The approach being taken to meet this challenge during SM-2 development involves placing a high priority on each of three major tasks:

1. Electronic Parts Improvement
2. Overstress Burn-In Testing
3. Reliability Growth Monitoring.

Each of these activities will be described below, together with some of the recent experiences at the Pomona facility that contributed to their selection and emphasis.

Electronic Parts Improvement. A major effort is under way to upgrade the reliability of SM-2 purchased electronic piece parts to levels significantly higher than those procured for SM-1. The reason for this action is that recent production experience on SM-1 has revealed that the major cause of test failure is the purchased electronic part.

The SM-1 production contract requires the demonstration of a very high success rate on a second acceptance test for each monthly lot of each of four missile sections. Two of the four sections pass quite easily, but the other two (more complex) sections fail the requirement quite frequently. As a consequence, it has been necessary to implement a policy of detailed diagnosis of each of the failures that occur during these success-rate demonstration tests.

An analysis of the results of these diagnoses (Figure 4) shows that 80 percent of the acceptance test failures involve electronic parts, that 73 percent of the failed parts are semi-conductors, that 75 percent of the part failures are supplier related, and that there are two primary causes of failure: (1) conductive particle contamination, and (2) defective internal bonds. Immediate corrective action was undertaken on the SM-1 program, including the scrapping of suspect lots and the added screening of the more offensive part numbers, using shock, centrifuge acceleration, temperature cycling, and loose-particle detection tests. In addition and in parallel, a larger-scale attack on the problem was initiated for SM-2.

During SM-2 development, a new set of specifications will be invoked for a higher grade of parts than had been specified for SM-1 (Figure 5). MIL-M-38510 (Class B), JANTXV, and ER (Level R) specifications will be employed wherever available. These basic standards will generally upgrade SM-2 parts reliability with regard to all of the various failure modes that occur. In addition, a part configuration (or "fingerprint") control will be imposed on the supplier to assure that significant changes in chip size, geometry, connections, sealing, etc., will be carefully evaluated and approved in advance. Finally, additional testing will be performed on semiconductors to assure detection of internal contamination and defective bonds.

These new controls, as they are being developed for SM-2 production, will also be phased into the purchase of parts for the development hardware, so that the reliability of the flight test rounds may benefit, and so that the tests and other controls can be proofed in advance of production. A continuing economic review of this approach will be conducted because of the potential impact of these improvements on parts costs, production repair costs, and the costs of ownership. The underlying objective will be to find the set of specifications, tests, and controls that provide the required production reliability for an acceptable production cost.[1]

Overstress Burn-In Testing. In addition to upgrading the electronic parts in the SM-2, a comprehensive burn-in test will be imposed on the development missiles. This test is designed to reveal the major failure modes inherent in the design so that they can be corrected prior to production. The stress levels will be higher than specification requirements, so that design safety margins will be checked and, in effect, an accelerated life test will be accomplished. As secondary benefits, the probability of success of each development flight test will be improved, and also the burn-in procedure will be proofed for subsequent use in production.

The usefulness of burn-in as a way of reducing infant mortality in missile assemblies has been demonstrated in the SM-1 production program (Figure 6). Missile sections are subjected to a series of vibration acceptance tests, with a minimum of three tests required for sell-off. When sections fail these tests, the next lower assemblies, called plates, are removed and repaired. When the average plate removal rate is plotted against the test sequence number, the reduction in infant mortality is clearly evident, with the region of constant failure rate being achieved at about the fourth or fifth test.

In an attempt to lower this failure rate during the SM-1 production program, an experiment was performed to evaluate several burn-in techniques on plates, prior to their assembly into sections. Plates were selected for this burn-in rather than modules or sections, as a compromise between the cost and the effectiveness of the test. As a result of the experiment, 100 percent burn-in of plates was introduced into production, employing a sequence of 10 hours of operation (1-hour on, 1-hour off), followed by 1 hour of monitored vibration with a 90-degree rotation every 15 minutes. The result of section-level acceptance testing was that the infant mortality curve was pushed to the left and downward (Figure 6). That is, infant mortality was still in evidence, but it was lower to begin with and then dropped down to a lower level of constant failure rate. In addition to a lower delivered failure rate, the plate burn-in caused a reduction in the average number of section tests required to achieve acceptance.[2]

Concurrently in the same production facility, the Standard ARM (Anti-Radiation Missile) production program also had very good results with a slightly different kind of burn-in. One-hundred percent of the Standard ARM packages (roughly equivalent to SM-1 plates) were subjected to a burn-in sequence of vibration, low temperature, high temperature, and repeat vibration, all while operating. The failure rates at higher assembly levels were then compared with those of the previous production run which had not employed this package-level burn-in. The section-level failure rate was down 60 percent for the guidance section and 80 percent for the autopilot section, and the missile-level failure rate was reduced by 40 percent. Discussions with other missile contractors have indicated that this kind of improvement is quite typical.[3]

As a result of these favorable results with burn-in testing, it was decided to adapt the concept to the SM-2 development program. A design margin (i.e., overstress) test policy had already been incorporated into the design process, however, a formal test was needed for the upper levels of assembly. A reliability evaluation test that would triple the test operating hours was considered, but it proved to be too costly and difficult to schedule. Overstress burn-in appeared to provide a good way to combine the benefits of all these different kinds of testing at a reasonable cost.

The objectives of the test (Figure 7) will be to: (1) provide failure mode data for design improvement, (2) remove infant mortality from the development hardware prior to flight test, and (3) prove the effectiveness of various burn-in stresses for subsequent use in the production program.

The candidate environmental stresses for this test include high and low temperature, vibration, and shock. The 10-hour test was taken from the SM-1 production burn-in with high temperature added to accelerate failures. The low-temperature test was taken from the Standard ARM experience cited above, which indicated low temperature to be a good way to detect defective connections. The vibration test was taken from the SM-1 burn-in, with the added requirement to shake in three planes. Shock was added to provide still another test of interconnections. A minimum running time objective has also been established so as to assure adequate data for MTBF measurement. The environmental stress levels are set at 1.5 times the specification requirements, i.e., 50 percent beyond the specification limit, referenced to the ambient level. This safety margin value was extensively employed during the development testing of the very successful REDEYE missile, and has also been set as the objective for SM-2 design capability.

Reliability Growth Monitoring. The two tasks described above will (1) address the biggest known problem (parts), and (2) provide a way of uncovering new and unknown problems (by testing). The third task, reliability growth monitoring, is designed to measure progress toward the reliability goal. The procedure here will be to plot a growth curve of SM-2 MTBF that can be followed during development to see if the final objective is likely to be achieved and to provide the impetus for additional action if it appears that the goal might not be met.

Guided missile reliability growth measurement can be a problem. If only 10 or 20 missiles are allotted for flight testing, it is difficult to get a precise point estimate of reliability for the whole sample, not to mention intermediate points on a growth curve. Pre-flight ground test success rates are also hard to score because the tests vary from one to another and from missile to missile. Operating time and MTBF measurements have traditionally been dismissed as being inappropriate to "one-shot" devices. But there is a compelling need for monitoring reliability accomplishment, and the MTBF parameter will be reconsidered for SM-2. It provides a convenient method for pooling all failures from all types of tests into a single index of growth.

Recent experience at Pomona with MTBF growth measurement on another system has been gratifying. That is, the MTBF growth plot has been easy to understand, easy to compute, easy to explain, and therefore has provided a useful management tool (Figure 8). Phalanx is a radar-controlled gun system now being developed by General Dynamics for the Navy. A graph of the measured Phalanx MTBF through each phase of its testing clearly shows growth toward the contract target. When plotted on log-log paper against cumulative operating time, the growth is approximately linear and therefore easy to extrapolate.[4] This technique will be applied to SM-2.

All pre-flight testing from the plate level up to the missile level will be subject to MTBF monitoring (Figure 9). Small operating time meters will be adhesively attached to each plate, and subsequently to each assemblied section. Test failures will be documented by an existing system for difficulty reporting. MTBF's will be computed at the conclusion of ground testing on each missile.

An MTBF goal will be established for SM-2, based upon the specification flight reliability goal, average predicted flight time, and an assumed exponential formula. MTBF measurements will be plotted on a log-log graph against cumulative operating time, with one point for each missile. An "adjusted" MTBF will also be plotted so as to discount failures that have been precluded from recurrence by redesign. This growth curve will be provided to upper management so that the improvement/redesign process can be redirected as necessary to achieve the goal.

Summary

Standard Missile 2 is challenged by a high reliability requirement and a need to add new functions so as to meet the threat. To meet this challenge, the reliability program is concentrating heavily on three areas of activity. The first task is to upgrade the purchased electronic parts, which are known to be a major cause of failures. The second task will impose a tough overstress burn-in test on development missiles so as to identify modes of failure in the new design. The third task will monitor MTBF growth and thereby measure progress toward the goal. Economic pressures have forced the SM-2 approach to reliability to be simple and direct, and recent experiences on other programs have indicated the logical steps to take.

References

1. Goldshine, G. D., and Martin, R. E., "Component Defects and System Reliability", 1973 Annual Symposium on Reliability, Philadelphia, Pennsylvania.

2. Macy, H. C., and Bullington, H. E., "Repeatability Assessment Program Final Report", 6 July 1971, General Dynamics, Pomona Division, Pomona, California.

3. Finley, J. E., and Haigh, W. B., "Stress Testing - The Answer to Reliability in the 70's!", 27 April 1972, IEEE Spring Symposium.

4. Shelby, J. D., and Miller, S. G., "Reliability Planning and Management (RPM)", 26 September 1970, ASQC/SRC Seminar, Niagara Falls, New York.

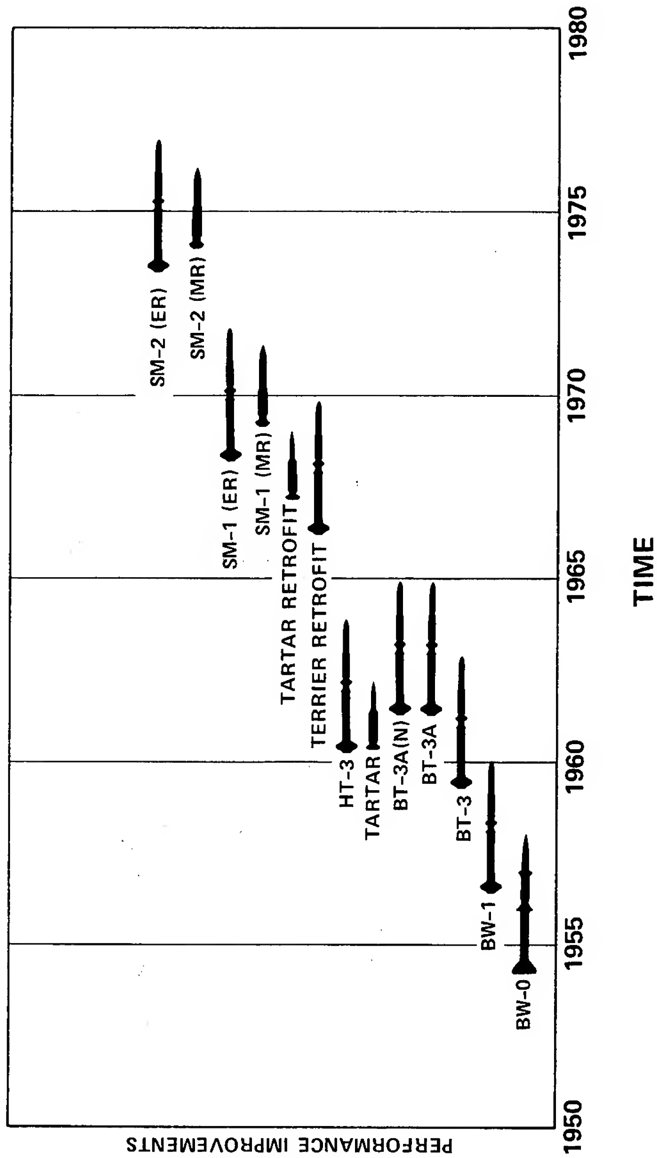


Figure 1. TERRIER/TARTAR/Standard Missile Evolution

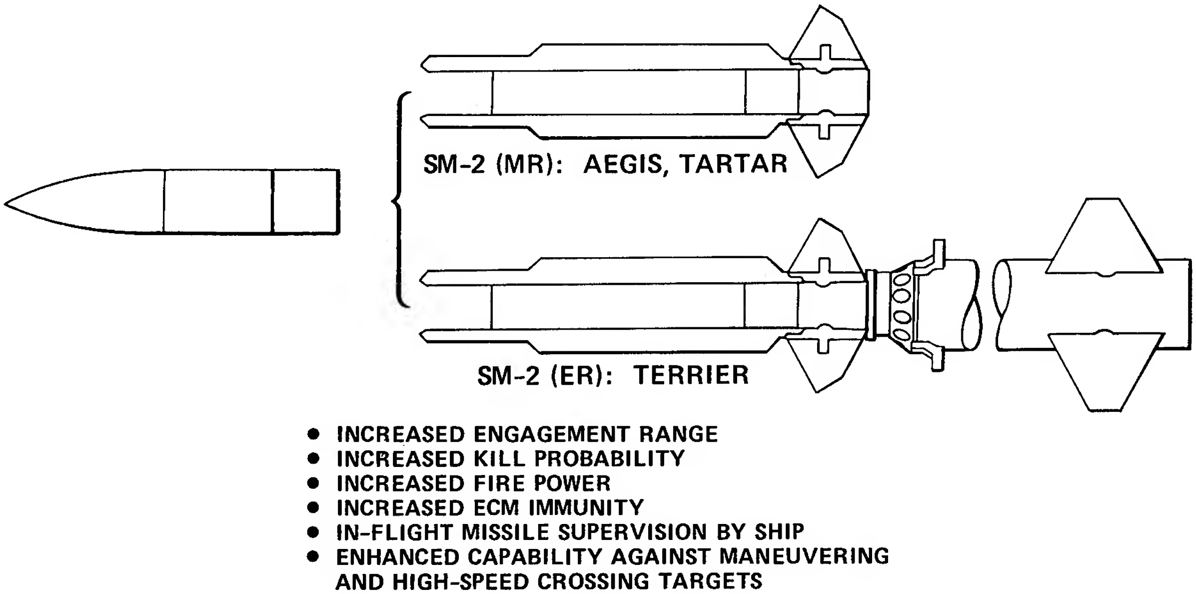


Figure 2. The SM-2 Missile

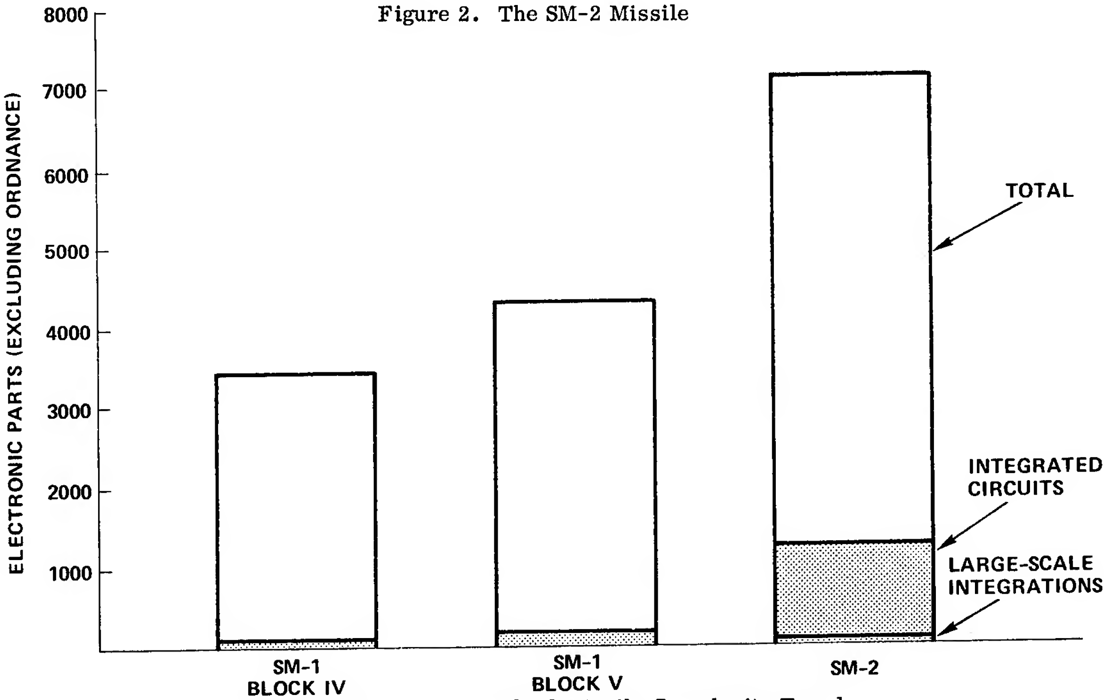


Figure 3. Standard Missile Complexity Trend

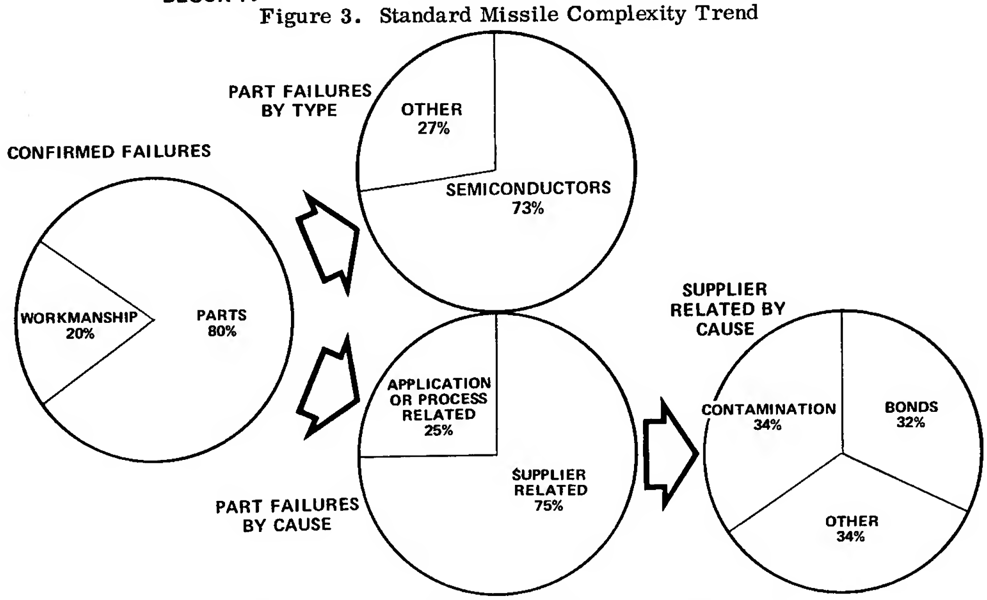


Figure 4. SM-1 Final Assembly Failure Diagnosis

SPECIFICATION LEVELS

- MICROCIRCUITS – MIL-M-38510, CLASS B
- TRANSISTORS AND DIODES – MIL-S-19500, JANTXV
- PASSIVE PARTS – VARIOUS MIL SPECS, ER LEVEL R

PART CONFIGURATION CONTROL

- ESTABLISH BASELINE ON CHIP SIZE, GEOMETRY, CONNECTIONS, SEALING, ETC.
- PERMISSION REQUIRED FOR SIGNIFICANT CHANGES
- SAMPLING INSPECTION OF EACH LOT

ADDED TESTING OF SEMICONDUCTORS

- BOND STRENGTH – SAMPLE TESTING AT SUPPLIER AND UPON RECEIPT
- LOOSE PARTICLE DETECTION (LPD) – SAMPLE TESTING AT SUPPLIER AND UPON RECEIPT, WITH 100% SCREENING OF FAILED LOTS
- TEMPERATURE CYCLING – 100% SCREENING UPON RECEIPT
- 100% ELECTRICAL TEST AFTER LPD AND TEMPERATURE CYCLING

Figure 5. Electronic Parts Improvement

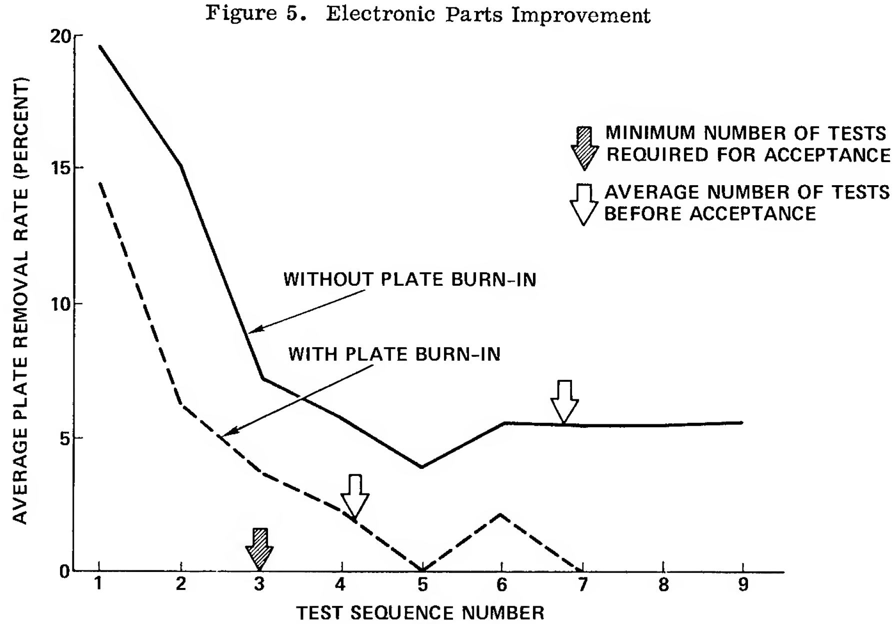


Figure 6. Infant Mortality in Standard Missile Sections

OBJECTIVES

- PROVIDE FAILURE MODE DATA FOR DESIGN IMPROVEMENT
- REMOVE INFANT MORTALITY PRIOR TO FLIGHT TEST
- PROOF BURN-IN PROCEDURES FOR SUBSEQUENT PRODUCTION

CANDIDATE STRESS

- HIGH TEMPERATURE – 10 HOURS AT HIGH-TEMPERATURE OVERSTRESS, WITH 5-MINUTE TEST EVERY HOUR
- LOW TEMPERATURE – 1 TEST AT LOW-TEMPERATURE OVERSTRESS
- VIBRATION – 20 MINUTES AT VIBRATION OVERSTRESS, WITH TEST IN EACH OF 3 PLANES
- SHOCK – 1 TEST AT SHOCK OVERSTRESS
- RUNNING TIME – 20 HOURS TOTAL BEFORE MISSILE ASSEMBLY

Figure 7. Overstress Burn-in of Development Missile Sections

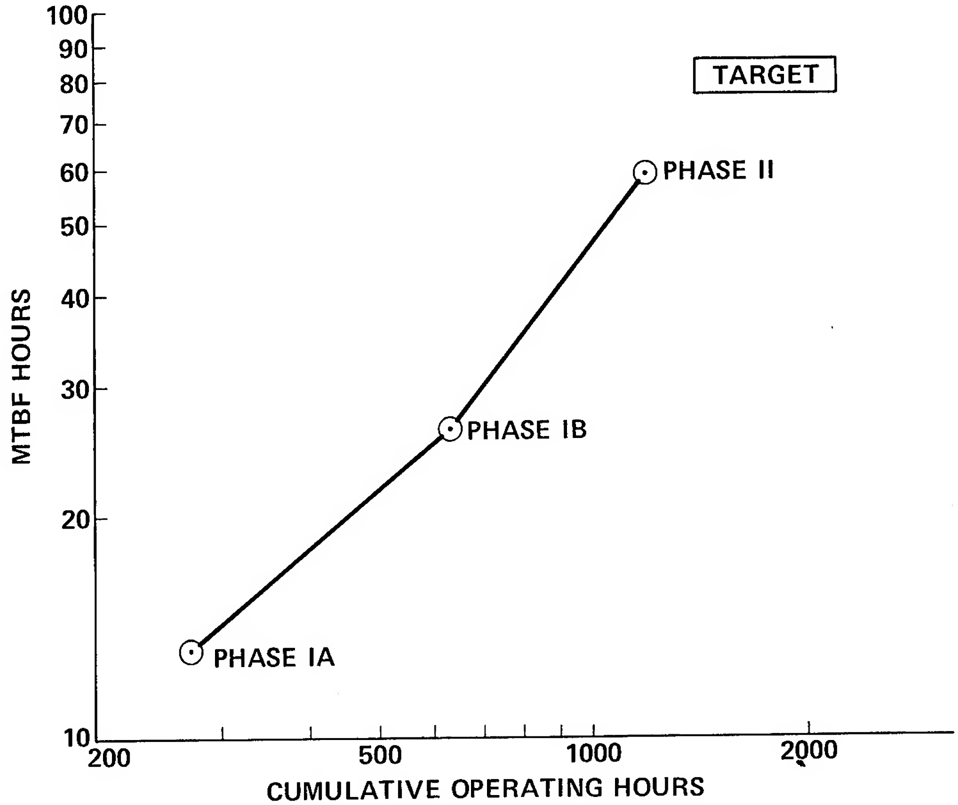


Figure 8. Phalanx System Reliability Growth

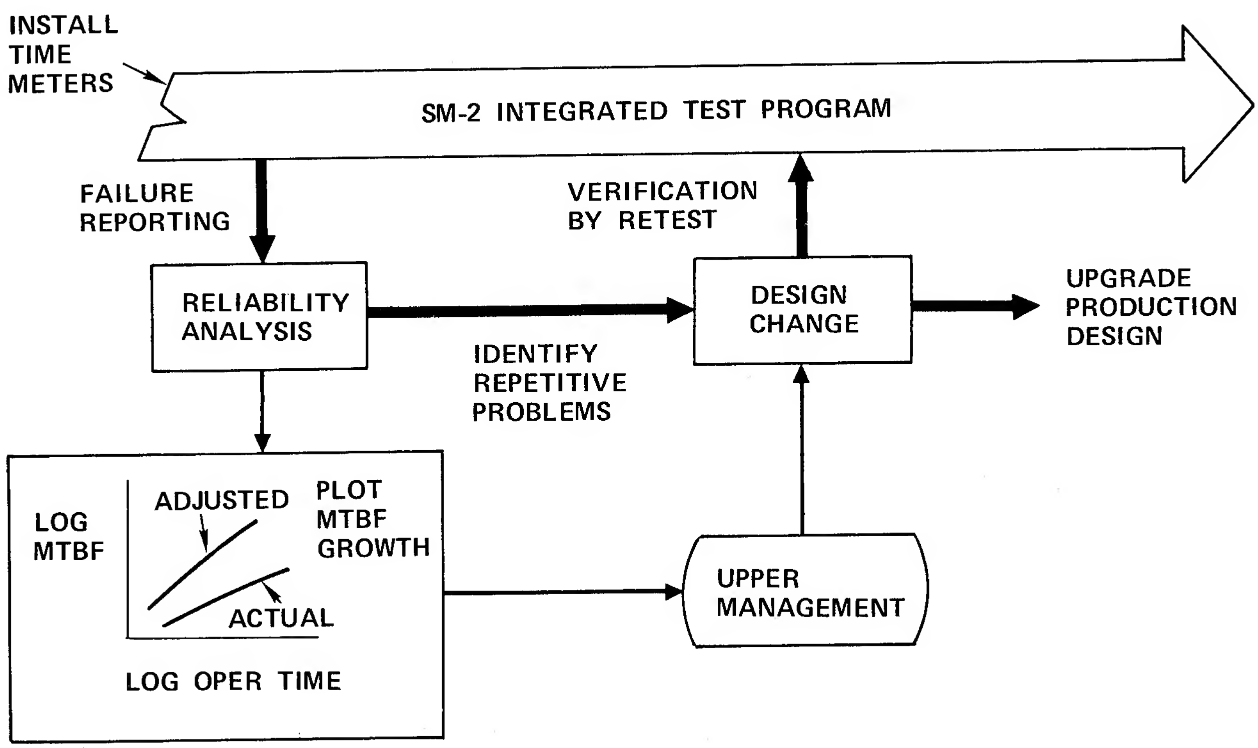


Figure 9. MTBF Growth Monitoring

RELIABILITY IN HOSPITAL INSTRUMENTATION

Karen Steel Kagey, M.D.
Peter Bent Brigham Hospital
Boston, Massachusetts

The importance of reliability in hospital instrumentation is often considered confined to life support devices such as heart-lung machines, or to be measured only in terms of fatal or near fatal incidents. This emphasis on the importance of specific appliances has tended to obscure other areas in which the failure of one piece to perform may lead to a string of other marginal performance or failure situations of less impressive nature, but which also detract from the smooth running of the hospital. Examples will be given of some of the many forms in which 'unreliability' may arise.

Application Problems

Equipment may appear unreliable because performance is as designed, but not as expected. Several times each year I am told that a demand-type cardiac pacemaker has failed to recognize spontaneous cardiac electrical activity, a situation which can lead to serious arrhythmias. In a large percentage of these cases the pacemaker is functioning perfectly, but the *system* has failed because the placement of the wires in contact with the heart resulted in an electrical signal from the patient which was below the threshold of the pacemaker sensing circuit. Continuous reiteration and training of staff is needed to overcome these problems, to make them aware a threshold exists and how to measure it. Sensing threshold figures are usually given somewhere in the literature about the device; but rarely appear on the instrument.

Control of body temperature when extreme fever occurs may be accomplished by cooling the skin to increase heat loss. This may be done with wet towels or with a special blanket through which fluid of a selected temperature is circulated. Since the hypothermia blanket is a machine, it is commonly assumed that it must be far more effective than old-fashioned wet methods. Nevertheless, the same principles apply to both methods, and only if the physiology is understood (and a method based on these principles followed) will the end result be as desired. Heat transfer occurs only where skin and cooling medium are in contact. Contact alone is not adequate. If the cooling medium is too cool, the blood flow to the skin is reduced by reflexes thus minimizing heat loss. Also, the patient may begin to shiver which substantially raises heat production, further compounding the problem.

The standard electrocardiograph responds to frequencies between perhaps 0.1Hz and 70Hz or higher. This is what cardiologists are accustomed to seeing. Some oscilloscopic ECG monitors have narrower bandwidths to eliminate artifacts caused by poorly designed electrodes, cables and poor electrode application. The effects of bandwidth narrowing vary from negligible to great; and, when large, usually simulate the pattern seen when blood flow to the heart is seriously compromised. Not many physicians have been educated along these lines, and it may be difficult from looking at a monitor front panel to know what the manufacturer designed it to do. Education of those responsible for use of equipment to obtain properly designed monitoring cables and apply them properly; education of the physicians to what potential 'information errors' may exist; adequate front panel information, will improve this situation.

The effectiveness of ultraviolet light (253.7nm) for sterilizing airborn bacteria is well established, and it is generally understood that some minimum power output is required; but the various factors which may influence such output seems less clear. Emphasis is placed on frequent cleaning of the lamps to remove accumulated dust or grease, and rather less emphasis on measurements to find out whether cleaning is needed. Little information is readily available on how reliable measurements are (there is said to be variation in NBS standards as well); what variability may be expected, lamp-to-lamp or ballast-to-ballast; what effect type of fixture, etc. may have in time from lighting to achievement of steady state output conditions. The last is an interesting phenomenon noted by one of our electricians who found that after relighting certain fixtures the output climbed for a period and then stabilized by the following day. It is probable that some of the pressure for frequent cleaning stems from early studies in which an increase in output after cleaning was attributed to cleaning, but was actually due to this phenomenon.

Humidity also is important for above about 65% relative humidity the effectiveness of UV radiation in killing bacteria plummets. In sum, ultraviolet radiation is a very effective sterilizing tool when used correctly; those planning an installation must look well beyond the initial installation in order to maintain effectiveness.

Environmental Problems

Performance failure may result from power line voltage fluctuations beyond the limits permissible for a particular device. The first problem is to obtain the limits from the manufacturer in evaluating a new purchase. They are rarely given in medical-equipment specification sheets. When these limits are likely to be exceeded in a given installation it is important to know what the effects will be: minor change in a monitor sweep speed; stalling of the motor of a ventilator; incorrect results from a blood analyzer; potential instrument damage. The reliability of performance in a given area is not so much the characteristics of the instrument; but knowledgeable planning to compensate for the instrument if interference with normal operation is to be expected.

Supplemental oxygen from a central or local compressed gas source may be vital to care of a patient, although the compressed gas source is not required per se for operation of a ventilator or humidifier. There are still such devices which use tapered nipple junctions and light weight plastic tubing to link the therapeutic device with the source. These delivery methods are prone to blockage of tubing by bed wheels and unperceived disconnection of the tubing from the nipple by activity in the area. The solution is simple: utilize standard threaded or other secure junctions and the more rugged high pressure tubing used commonly with other inhalation therapy equipment. When flow meters are required, mount them on the end-use device with high pressure tubing between flow meter and source. This last also avoids having several flow

meters ganged on Y's into one outlet, a bunching which sets the stage for mistaking the flow meter of one device for another when making flow alterations.

A perhaps inadequately understood fact-of-life in hospitals is that even with maximum staff training and interest, when patient emergencies arise all attention is focused on that situation with resultant occasional maltreatment of objects in the area. And maximum staff cooperation in non-emergency situations seems an elusive object at best. The following two examples indicate ways in which these facts enter the reliability picture.

When central suction systems are installed without screens or traps at each inlet, successive spills of proteinaceous fluids build up deposits along the pipes. Efficiency is gradually reduced and the suction to a large or small area is finally lost until repiping is accomplished. When screens are used at each outlet, one still has the problem of when to change them. Suctioning is almost always performed with a partly, not completely filled catheter, thus some dynamic measurement of system effectiveness, not simply the maximum vacuum drawn over a long test period. Such a clinically oriented testing device would assist materially in scheduling maintenance when the area is free and when personnel are available. One problem with maintenance in hospitals is that the ever-present spector of true emergencies casts a 'life-or-death' glow over almost everything with resultant inefficiencies in running an engineering department.

Another class of victim is the electric/electronic appliance with perforated enclosures which permit entrance of spilled or splashed liquids. Since these liquids are often salt solutions, serious internal damage may result taking the instrument out of service (perhaps at a vital time) until repairs are effected. Unperforated horizontal surfaces, splash-proof louvers, gasketing of control openings, and retro-fitting old equipment with some sort of covering will help reduce this mode of equipment failure to a minimum.

A new and different instance of 'environmental pollution' occurred recently involving a heart-lung machine which ran-away in the automatic pumping mode during an open-heart procedure. Extensive testing of the machine failed to reveal internal malfunction, and after several weeks of mystification it was noted that the problem occurred only in one operating room, and then only when the large light in that room was somewhere between full on and full off intensity. The problem was found to be signals injected into the room power wiring from the SCR dimming circuits of the lamp. These were interpreted by the pump control as signals from its own speed circuits. The light manufacturer had not considered what the impact of his light control circuits would be on an isolated power distribution system, and the heart-lung machine manufacturer had not considered the possibility of such interference. Solutions are still being evaluated.

Failure of electric bed controls may cause nuisance situations, or may endanger resuscitation efforts if the head cannot be lowered in an emergency situation. Spontaneous activitation of bed controls can also present hazards. Apparently one manufacturer who incorporated optical switching in the hand control to isolate the patient from powered circuits failed to enclose the controls in a suitably rugged case. Cracks in the housing apparently permitted ambient light to enter and trigger the 'spontaneous' movements.

<u>Power Source Availability</u>

Line electrical power is assumed to be constantly available and infinite in capacity. Line operated devices are multiplying at rapid rates, not only for special monitoring and therapeutic needs, but for comfort, convenience and entertainment. The last may seem unimportant but actually for many patients in all degrees of illness it is an important adjunct in care. As more items are brought into the patient care area more outlets are required - especially where leakage current considerations dictate short cords which rapidly become trip hazards if outlets are not near the device. Most of us work within existing, much modified institutions, and unless great pains have been taken with every electrical renovation branch circuit diagrams may be very unreliable. Administrators must be made aware of the need to put the time and manpower into assessment of current resources, and cooperation between medical and engineering staff is necessary to plan realistically for the future. Additional outlets added on existing branch circuits will increase the probability of the circuit opening because of overload, especially dangerous if outlets on this circuit are scattered through several patient areas, or if total power to one room is lost. Circuit distribution should be planned for the specific area, including consideration of the needs of food trucks, housekeeping equipment and emergency equipment. The medical staff needs education here also, so that if limitations on equipment to be used in specific areas is necessary, the reasons will be understood and followed.

Power availability in event of normal-source failure is a separate but related question. Total coverage with emergency generators is appealing, but if this is not realistic for a given institution careful appraisal of true emergency needs is necessary if adequate provisions are to be made. In an intensive care unit one should have emergency power accessible at each bedside since it is not possible to predict when outages may occur, and thus place patients requiring emergency power near one or two special outlets. Each bed, ideally, would be on an individual circuit breaker so that problems at one bed will not affect other patients. However, for this system to function, the staff of the unit must be educated as to what equipment is emergency equipment, and how far the system can be loaded. This has seemed to work in our hospital ICU-Recovery Room but only with a written procedure and periodic discussions. It cannot be overstressed that education is important and must be ongoing. During one blackout I visited our artificial kidney unit and discovered that 5 or 10 minutes had elapsed before someone 'discovered' the emergency power outlets. When I arrived the prime concern was 'how long the batteries would last'. Explanation that a generator, rather than batteries, was the source of power proved reassuring to all.

Batteries, themselves, sometimes present problems in availability. In the past there has been a tendency for some instruments to be designed for special batteries obtainable only through the instrument manufacturer. This can lead to feast-and-famine situations such as happened several years ago at our hospital. Three different areas all had one or more units of a specific instrument. All ran out of special batteries at once and each area ordered substantial quantities. These quantities could not have been used up over two or three times shelf life, thus represented much wasted money. We have since consolidated stocks to minimize overstocking, but it is still easier to keep up with the instruments which use batteries available at local supply stores.

Batteries have figured recently in discussions of 'uninterruptible' power systems which utilize battery-inverters to bridge the gap between power loss and

restoration by other means. There are actually very few items for which a five to ten second outage represents serious hazards except for computers. A small instrument which must function continuously can be equipped with a suitable battery pack which has the advantage of permitting patient transportation from one area to another without problem. Battery systems capable of holding a 90-100 ampere load for a useful period of time will be costly and require careful maintenance if they are to be reliable. On the other hand, careful maintenance of the emergency diesel or other generator with regular testing of time from power loss to restoration under load will fill almost all needs. If the generator does not start at all, the battery-inverter systems probably will not hold enough equipment to be extremely useful. However, such generator load testing takes manpower, in our case overtime since load tests are run on Saturday when potential failure (should it occur) will be least hazardous to patients, and again medical staff and administrators must recognize the need for these expenditures. We also record from the line during the run to have an accurate record of time from power loss to restoration. This has two uses: first, we know we should expect power restoration within 5 seconds; second, when about 8 seconds was required during one test it warned that something was amiss and this was found and corrected before it became a serious hazard.

Hazard Generation

A medical device is expected to be safe to patients and personnel, and can be considered unreliable if it becomes potentially hazardous. Some devices such as X-Ray machines and defibrillators are inherently dangerous and safety involves separating persons from all unnecessary contact with known hazardous outputs. The needs here are usually well recognized. Another active-output device, rarely recognized as such, is the common pressure amplifier. The excitation current supplied to drive the transducer is prevented from reaching the patient by various insulation or isolation techniques in the transducer. However, if this excitation current is applied directly to the patient harm may occur. Burns were noted at the left arm and left leg sites of application of ECG needle monitoring electrodes during a surgical procedure. The ECG amplifier was immediately adjudged the culprit, but prompt asking of the right questions revealed that the ECG cable had indeed been connected to the pressure amplifier for a few moments. The problem was that identical connectors were used for the active pressure amplifier, and the passive ECG amplifier; and where a cable can be connected to the wrong place, it will be occasionally. We subsequently replaced the ECG amplifier connectors with a different type so that this hazard should be eliminated. However, at least two manufacturers still supply equipment where this can happen.

In the medical instrumentation field a by-word these days is 'leakage current' and what to do about it. It is recognized by many that the commonly used 18AWG cord sets with molded plugs are unreliable from the point of maintaining grounding wire continuity with the plug U-ground pin. Several of my nursing staff will attest to this, having been surprised by tingles from several ultrasonic nebulizers of older design with about 800 microamperes on the frame when the ground opens. We and many others have responded by using more rugged hand-wired plugs wherever possible, and insisting on them for new equipment. A well-designed molded plug and cord set could well be even better; but one has not yet appeared. Additionally, if one does appear it is unclear how it would be evaluated. There are apparently inadequate standards for grounding pin or receptacle grounding outlet reliability at this time. In addition to making choice of purchase difficult, this lack of standards, and resultant inadequate components for hospital misuse, has lead to consideration of totally different configurations for hospital use. If it is demonstrated that a parallel-blade U-ground configuration cannot be produced in reliable forms for hospital use, then other alternatives may be necessary. However, radical departure from current components necessitates expensive remodeling and usually 'adapters' to convert one power configuration to the other - such adapters producing major problems in supply, maintenance, etc.; and to go to such lengths if the reliability of well-designed conventional configurations could be adequate seems less than ideal.

On the other side of the leakage current coin, a reliable instrument should not develop large leakage currents under normal operation. When we first started bringing kidney dialysis machines to the Intensive Care Unit, which happens to have an isolated power system with ground fault alarm, we found that alarms were frequent. Inspection of all eight machines revealed that three carried almost line voltage on the frame during use, one carried about 5 volts, and the others were within reasonable limits. The fault was traced to the type of heaters and thermoswitches used in the salt bath, neither of which were designed to operate in this environment. Replacement with components capable of operating in saline has eliminated gross problems although we are continuing bi-weekly testing at this time. Here the problem was not noticed by the staff since operation of the machines was not impaired, and the fact that they were equipped with #12AWG heavy duty cords and good plugs maintained adequate frame grounding to eliminate shock possibilities.

Adaptability

To be reliable, an instrument must be usable when needed. Fiberoptic diagnostic instruments are becoming increasingly popular, and have allowed considerable advances in many areas. Certain companies specialize in specific types of instruments and thus a hospital may have instruments from several manufacturers for a variety of procedures. Each instrument must have a light-source, and interchangeability here leaves much to be desired. One manufacturer makes several adapters for use of other instruments with his light source but in other cases adapters are not provided, nor is the information which might enable a hospital to have a local machinist make one up. Often one is told that 'the other manufacturer's light is inadequate for my instrument', but at two o'clock in the morning bronchoscopy to remove a peanut from a lung can be more readily carried out with sub-optimum light than with no light.

Another example of inflexibility of routine equipment is the conventional Nurse Call system by which the patient activates a buzzer by pressing a small button with his thumb. This presents no problem until we get to the patient without hands, with heavily bandaged hands or who is paralyzed. If he is also unable to speak loudly, if at all, he may experience frightening and frustrating inability to contact floor personnel. It cannot be too strongly emphasized that the paralyzed patient, the patient being ventilated by a machine, the badly injured patient, may be just as alert as a healthy individual and these communication problems are real. We have just begun experimenting with devices which can be activated by foot, shoulder, or other existing controllable movement. It apparently has not been done commercially, and it only took me four years

to think of doing it now. The volume of this need is small, monetarily probably minimal, but when it is needed it extends whatever reliability the Nurse Call system has for communication to patients not previously able to use it.

It is always easy to list faults, and in the medical field current movements are prone to jump on these as indicative of a very poor situation. Actually, much good is done in hospitals, but more could be done. Reliability involves close involvement of supplier and user, and realistic approaches to hospital care needs. There will always be trade-offs between complexity, expense, size, maintainability, etc. The key is proper selection of instrumentation for carefully assessed needs, and then proper and continuing education of the users.

RELIABILITY AND PERFORMANCE CRITERIA FOR ELECTRO-MEDICAL APPARATUS
A Hospital Engineer's Viewpoint

INDEX SERIAL NUMBER — 1051

Bernard Stein, C.B.M.E.T.
Charles S. Wilson Memorial Hospital
Johnson City, New York

When an instrument is used in a hospital for patient care, it is subject to a large variety of conditions which impose unique situations that would not be encountered elsewhere.

1. The equipment is subject to misuse.
2. The equipment is usually operated by personnel who are not familiar with the full range of capabilities of the instrument.
3. The equipment is either never or poorly maintained.

This imposes a severe problem for the conscientious manufacturer and his engineering team who must develop this instrumentation to operate under these conditions.

Very simply stated, the reliability and performance criteria can be described by three classes:

Class 1 - Those instruments that upon malfunction or misuse are capable of inducing an injury to either the patient or the operator.

Class 2 - Those instruments that upon malfunction or misuse are capable of inducing a lethal injury to the patient or operator.

Class 3 - Those instruments that upon malfunction or misuse are incapable of inducing injury, lethal or sub-lethal to the patient or operator.

Examples of Class 1 instrumentation in use today are the typical electric beds used in a hospital or a suction pump which is ungrounded

Class 2 instrumentation is best illustrated by line-powered cardiac pacemakers and cardiac defibrillators.

Class 3 instrumentation is best illustrated by the latest patient monitoring equipment that uses optical isolation techniques between the patient and the electronics.

Now that I have described the three classifications for performance and equipment failure, I will now describe the classes of use for the electro-medical apparatus. The first use is equipment that is used in "Electrically Sensitive Areas." This equipment could best be described as follows:

> Equipment intended to be connected to an electrically conducting path onto the skin surface or into the blood vessels of the patient (e.g. by surface electrodes, needle electrodes, catheters), or intended to be used in a procedure which would result in a conducting path to the low internal impedance of the patient's body, either through a surgically created opening or through one of the natural orifices. This definition includes a conducting path formed by a fluid column such as that used with a dialysis machine or a blood pressure transducer.

An example of a typical electrically invasive technique is the use of an intra-cardiac electrode for taking a "V" lead electrocardiagram from inside the heart to a typical EKG machine which in this case, could well become a Class 2 type instrument. Another example is a fluid-filled catheter with saline as the conducting column for measuring intra-cardiac pressures. When this is connected to a typical pressure monitor, it too becomes a Class 2 type instrument.

The other type of equipment usage is known as equipment which is used in "Non-electrically Sensitive Areas." This type of equipment is described as:

> Equipment which is not intended to be connected to the patient by any electrically conducting means, other than casual contact with a grounded enclosure or frame.

Within the practical financial constraints, all electro-medical apparatus should be treated as Class 3 equipment since it may ultimately be used in "Electrically Sensitive Areas" of a hospital. Therefore, this equipment, even though it may be used in general care areas most of the time, must fall into a Class 3 reliability class, which is the type of instrument that upon malfunction or misuse is incapable of inducing injury, lethal or sub-lethal to the patient or the operator. Unfortunately, in the way the state of the art exists today, most equipment that is used in "Electrically Sensitive Areas" of a hospital falls into a Class 1 or Class 2 category. The ideal goal, therefore, is to design equipment which falls into the Class 3 category.

Another equally important facet of this picture is the serviceability of the equipment. The equipment should be designed so that the Bio-Medical Electronics Technician (B.M.E.T.) in the hospital can service the equipment with a minimum amount of test equipment. Parts should be standard parts readily available, and the equipment should be set-up in such a way that test points and other alignment points are readily accessible to the technician.

An example of this type of serviceability is all P.C. boards should be plug-in, with all calibration adjustments and test points located at the top of the board and plainly marked for the B.M.E.T. to see. All transistors and I.C.'s should be in plug-in sockets. Parts should be plainly marked on the board and the schematics should list commercial parts numbers instead of special manufacturers' code numbers. Panel lamps should be easily replaceable, as well as switches and other controls. Circuit breakers should be used in place of fuses whenever possible and located where one can reach them. A front panel lamp should be incorporated showing a breaker has tripped. These and a host of other features such as quality components, glass epoxy P.C. boards, plated through eyelets on the boards, good wiring practice, conservative design and double insulation (the use of high impact plastics versus metal) for all exposed surfaces. High quality wires and cable assemblies should be provided with the instrument, as well as a complete set of decent service and operating manuals. Good design also demands the patient to be isolated by as high an impedance as possible from the electronics (i.e. optical coupling).

In summation then, the goal we are striving for in equipment reliability and performance is equipment that falls into a Class 3 category, and equipment that should be designed with a B.M.E.T. in mind, as well as the fact that the equipment may find its way into an "Electrically Sensitive Area" in its normal lifetime.

QUANTIZED GOALS FOR THE DESIGN OF VIBRATION AND SHOCK TEST FIXTURES

by Wayne Tustin, Tustin Institute of Technology, Inc., Santa Barbara, California

ABSTRACT

Vibration and shock testing are useful tools for assessing the ruggedness of units whose reliability may be reduced by vibration and shock inputs during handling and shipment as well as in actual service. Reliability specialists should insist that careful attention be paid to many elements of vibration and shock testing, particularly to the fixture which attaches test units to shakers and to shock test machines. Unfortunately, in many test organizations, little attention is paid to fixture design and behavior, and many items are either overtested or undertested. This article offers workable goals for the dynamic behavior of fixtures and states these in such a manner that an experimental investigation will show whether the goals have been met.

SUMMARY

1. Poor fixtures can foul up vibration and shock tests.
2. Realistic fixture design goals are needed.
3. Table I suggests design goals.
4. Experimental verification should follow fabrication.
5. Design of fixtures is a specialty.

INTRODUCTION

This article concerns fixtures -- structures used as in Figure 1 to attach vibration test specimens to shakers. Specifications for the dynamic behavior of fixtures are suggested.

Fixtures plan an important part in the results of vibration tests, but test specification writers tend to ignore this. They sometimes fail to even mention the fixture. They may specify dynamic behavior that cannot be achieved, or their wording may be vague and subject to several interpretations. Two examples follow:

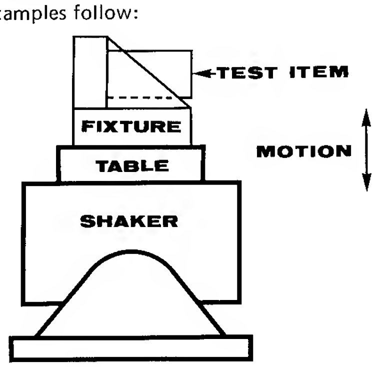


Figure 1 Fixtures are used to attach test items to shakers for vibration tests, also to shock test machines for shock tests.

MILITARY STANDARD 810 - B, Method 514.1, Paragraph 4.2, Mounting Techniques - *In accordance with Section 3, General Requirements, Paragraph 3.2.2, the test item shall be attached to the vibration exciter table by its normal mounting means or by means of a rigid fixture capable of transmitting the vibration conditions specified herein. Precautions shall be taken in the establishment of mechanical interfaces to minimize the introduction of undesireable responses in the test setup. Whenever possible, the test load shall be distributed uniformly on the vibration exciter table in order to minimize effects of unbalanced loads. Vibration amplitudes and frequencies shall be measured by techniques that will not significantly affect test item input control or response. The input control sensing device(s) shall be rigidly attached to the vibration table or to the intermediate structure, if used, at or as near as possible to the attachment point(s) of the test item.*

SANDIA CORPORATION STANDARD SC-4452D(M), page D 1.7, Section 2.3.4, Design Frequency - *Fixtures should be designed for as high a resonant frequency as possible. It must be realized that increasing the resonant frequency of a fixture is not always as simple as it first may appear. For example, to double the frequency of a simple spring mass system requires an increase of four in spring stiffness. The acceleration gradient across the height of the fixture is one problem frequently encountered with a vertical type mounting fixture. As a general rule, try to design a fixture to have a resonance three times the maximum test frequency and the acceleration gradient will be 10 per cent or less.*

Company documents are generally no more explicit. Consequently, fixture design is usually based on necessity for speed, availability of materials and fabrication capabilities, rather than on need for proper dynamic behavior.

Most specifications assume that fixtures can be considered "rigid". The goal of a resonance "three times the maximum test frequency" regardless of specimen size can seldom be achieved. In most practical cases, this is not possible and resonances usually occur during tests, particularly when the test item is attached. Table I summarizes practical, quantized guides or rules to guide the designer. Fixtures, once built, should be evaluated with loads that simulate test specimens, to demonstrate that goals have been achieved.

A BAD EXAMPLE

Poor test fixtures often contribute to a lack of repeatability between "identical" tests using different fixtures. Figures 2 and 3 show two different fixtures for testing a module which in service is cantilevered from a panel, attached by four bolts. The test goal was to identify the module's first two major resonances so that a "resonant dwell" test might be performed at each resonance. With the usual lack of quantized criteria, either fixture might be considered acceptable.

Based on the X_2/X_1 transmissibility graph of Figure 2, one might conclude that the specimen's only resonance was at 80 Hz. The fixture's first resonance was separately determined to be at 375 Hz, but there is no sign of it in Figure 2. Y-axis orthogonal motion was much greater than the desired X-axis motion at 80 Hz, but that is not shown by Figure 2. Nor is the fact that the 80 Hz resonance was not the lowest specimen resonance.

Later, the specimen was tested in a much stiffer fixture, as shown in Figure 3. The accompanying X_2/X_1 transmissibility graph shows a clean "classical" first peak at 40 Hz, also another resonant response at 82 Hz. This more rigid fixture eliminated interference between test item and fixture; there was very little Y-axis motion.

Suppose that laboratory A uses the improper setup of Figure 2 and conducts the subsequent life test only at 80 Hz. The specimen will probably pass. And suppose that laboratory B conducts its test with the mounting of Figure 3. Perhaps B interprets the graph such that two life tests are run, at 40 Hz and at 80 Hz. Perhaps the specimen fails. Which test is right? Each lab met its own interpretation of the test cri-

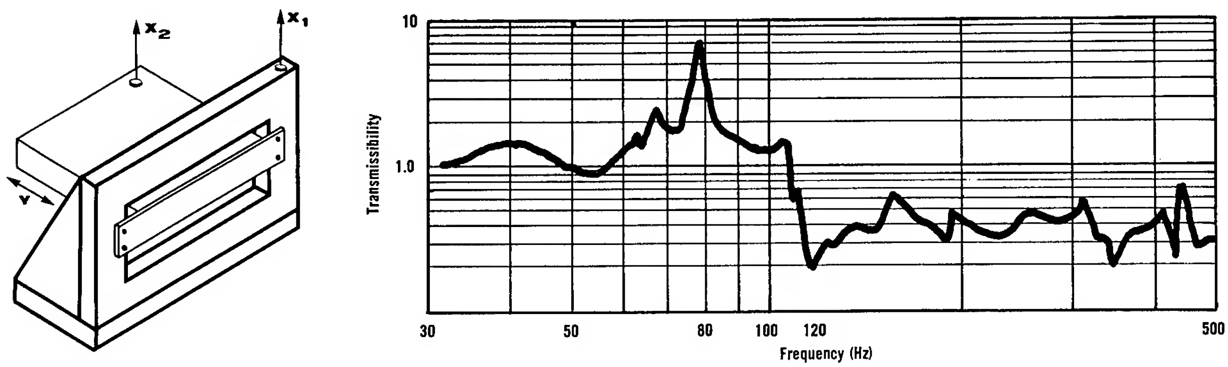


Figure 2 Results of a resonance search upon a typical electronic assembly, using an inadequate test fixture that flexes in the Y direction.

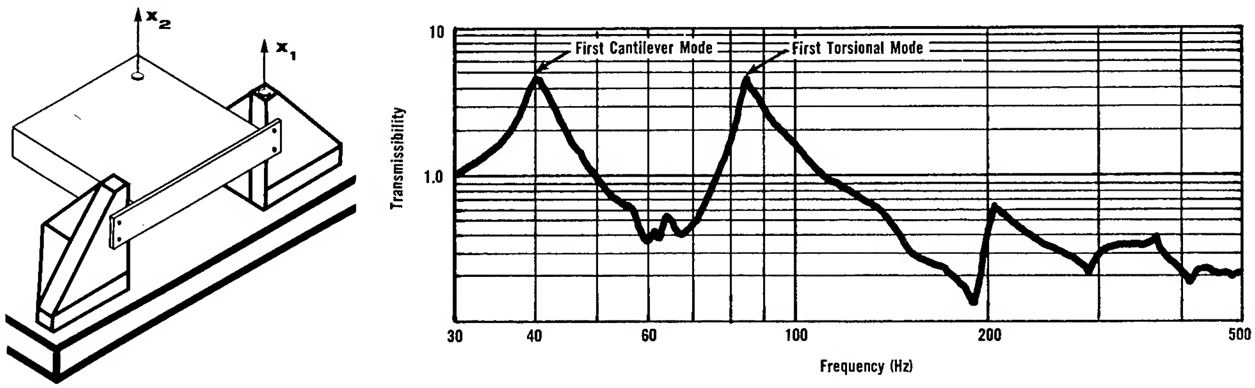


Figure 3 Results of a resonance search upon the same electronic assembly, using a much stiffer fixture with no flexing in the Y direction.

teria. Quantized fixture design criteria would have avoided these differences.

DISCUSSION

DESIGN CRITERIA CHART

How good must a fixture be? Table I presents some long-needed criteria or goals for the designer. It states the

1. Allowable transmissibility peaks.
2. Allowable amounts of orthogonal motion, and
3. Allowable variations in vibratory input between attachment points to the test item.

Fixtures designed for vibration testing are also usable for shock testing; thus the Chart of Table I applies to fixtures for both shock and vibratior testing. The Chart is proposed for inclusion in future Military and other Test Standards, as well as for the preparation of detailed test specifications.

Proving that the Chart has been complied with requires dynamic measurements of fixture behavior; reliability specialists should insist that test reports include those measurements. The fixture should be loaded with a dynamically similar "dummy" test item, a prototype or, best, the test item itself. The four columns of the Chart will now be discussed.

COMPONENT DESCRIPTION. Enter on the horizontal line which most nearly matches your test item size and weight.

ALLOWABLE TRANSMISSIBILITY PEAKS. This column places limits upon the number of resonant peaks in the fixture's response curve.

ALLOWABLE ORTHOGONAL MOTION. This column places limits upon the amount of lateral axis motion.

ALLOWABLE VARIATION IN MOTION INPUT. This column places limits upon the variations in motion intensity among the several specimen/fixture attachment points. Inputs to the specimen are generally equal at low test frequencies where no major fixture or specimen resonances occur. As test frequency rises and resonances occur, variations are caused by resonant/antiresonant responses in the moving system. One attachment point can respond to a resonance while another responds to an antiresonance, thus the motions of attachment points are greatly different.

The criteria of Table I can sometimes be bettered. There will be instances in which these criteria cannot be met. However, these goals are felt to be reasonable and generally attainable. The Chart may thus be used as a starting point for negotiations and design.

CRITERIA RELAXED FOR LARGE ITEMS

Not all tests go to 2,000 Hz. Typically, a 1,000 pound test object will not respond above 500 Hz; its internal parts will not respond at higher test frequencies applied through the attachments.

The column ALLOWABLE TRANSMISSIBILITY PEAKS is based upon specimen response considerations, as well as upon the fact that large, heavy items, in service, receive very little high frequency excitation. The ORTHOGONAL MOTION and the VARIATION IN MOTION columns are based upon the upper frequency expected in service plus the first allowable transmissibility peaks (second column). Authorities highly recommend that any article weighing more than 50 pounds not be tested to the full intensity, full frequency range of a test specification. A waiver should be requested in most such cases.

FIXTURE RESONANCES CAN BE PREDICTED

The various sources of difficulty that arise in designing fixtures to meet the criteria of Table I will be demonstrated on a typical fixture. They are more troublesome on large fixtures. One can seldom separate

TABLE 1 — DESIGN CRITERIA FOR VARIOUS SIZES OF FIXTURES

Component Description	Allowable Transmissibility Peaks	Allowable Orthogonal Motion	Allowable Variation in Vibratory Input between Test Item Attachment Points
Small components, mechanical electrical, or electronic, up to cigar-box size and weight up to 5 pounds.	None below 1000 Hz. Above 1000 Hz, a maximum of 3 resonances, limited to 5:1 over 3db bandwidth 100 Hz.	Y and Z motions less than X motion throughout the test range up to 2000 Hz.	± 20% allowable up to 1000 Hz. From 1000 Hz to 2000 Hz, ± 50%.
Electrical, electronic, mechanical components in sizes up to a 10-inch cube and weights up to 15 pounds.	None below 1000 Hz. Max. of 4 peaks above 1000 Hz, 5:1. None to exceed a 3db bandwidth of 100 Hz.	Y and Z motions less than X motion throughout the test range up to 2000 Hz.	± 30% up to 1000 Hz. 1000-2000 Hz, not exceed 2:1 between any pair of points.
Odd-shaped mechanical components (i.e., large hydraulic actuators and vent relief valves). Electrical equipment (i.e., inverters, telemetering transmiters). Volumes up to 3 ft^3, weights 10 to 50 pounds.	None below 800 Hz. Max. 4 peaks 6:1 over 3 db bandwidth 100 Hz, 800-1500 Hz. Max. 3 peaks 8:1 over 3 db bandwidth of 125 Hz, 1500-2000 Hz.	Y and Z motions less than X motion up to 1000 Hz. Above 1000 Hz, 2X, except that over a 3 db bandwidth of 200 Hz, may be 3X.	± 50% up to 1000 Hz. From 1000 Hz to 2000 Hz, 2:1, except that over a 3db bandwidth of 200 Hz, input variation may be 2.5:1 between any pair of points.
Larger equipment weighing 50 to 500 pounds, volumes up to 20 ft^3.	None below 500 Hz. Max. 2 peaks 6:1 over 3db bandwidth 125 Hz, 500-1000 Hz. Max. 3 peaks 8:1 over 3db bandwidth 150 Hz, 1000-2000 Hz.	Y and Z less than X to 500 Hz. 500-1000 Hz, less than 2 X, and 1000-2000 Hz, less than 2.5 X, except over a 3db bandwidth of 200 Hz, may be 3 X.	± 50% up to 500 Hz. From 500 Hz-1000 Hz, 2:1 and 1000 Hz-2000 Hz, 2.5:1 except over 3db bandwidth of 200 Hz, variation may be 3:1.
Large equipment over 500 pounds and 24 inches minimum dimension. Note: These fixtures are exceedingly difficult to design. In general, use only with auxiliary hydrostatic bearings.	None below 150 Hz. Max. 1 peak 3:1 150-300 Hz also max. 3 peaks 5:1 over 3db bandwidth 100 Hz, 300-1000 Hz. Max. 5 peaks 10:1 over 3db bandwidth 200 Hz, 1000-2000 Hz.	Y and Z less than 1.5X up to 300 Hz. Less than 2.5X, 250-2000 Hz except over 3db bandwidth of 100 Hz in range 300-1000 Hz, may be 3:1; also over 3db bandwidth of 150 Hz in range 1000-2000 Hz, may be 4:1. Vertical motion not to exceed 1.5X over entire test frequency range. Use hydrostatic bearings.	± 50% up to 400 Hz. From 400 Hz-2000 Hz, 2:1 except over 3db bandwidth of 200 Hz, variation between points may be 3:1. © Tustin Institute 1971

transmissibility peaks (in the desired test direction) from orthogonal motion; this was briefly discussed in connection with Figures 2 and 3. The effect of transmissibility peaks and orthogonal motion will be analyzed separately and will then be combined to obtain the total response. The natural frequencies to consider are

1) gusset beam bending

$$f_g = 3.13\sqrt{\frac{K}{W}}$$

where $K = \frac{Ebh^3}{6L^3}$

and $W = \frac{1}{2} b_1 L_1 h_1 \rho$

and/or beam bending

$$f_B = 3.13\sqrt{\frac{K}{W}}$$

where $K = \frac{B_1 EI}{L^3}$

and $W = b_1 L_1 h_1 \rho$

2) plate bending of those surfaces acting as plates

$$f_p = 92.5 \cdot 10^3 \lambda \left(\frac{h}{b^2}\right)$$

(λ is a plate shape constant.)

3) fixture rigid body rotation (involves attachment bolts.)

$$f_{RBR} = 3.13\sqrt{\frac{nK_r}{TL}}$$

and $K_r = \frac{4EI}{L_B}$

4) base plate twist or torsion

$$f_{BT} = 3.13\sqrt{\frac{K_r}{TL}}$$

where $K_r = \frac{2GJ}{L}$

and TL is the torque moment arm.

5) the total frequency f_T is found by proper summation of frequencies 1) through 4), according to Dunkerley's Equation

$$\frac{1}{f_T^2} = \frac{1}{f_g^2} + \frac{1}{f_p^2} + \frac{1}{f_{RBR}^2} + \frac{1}{f_{BT}^2},$$

usually aided by a nomograph in Reference 1.

Simply due to size, large fixtures have relatively low natural frequencies. In addition, large fixtures often have many elements which combine to a lower f_T than any of the individual elements. These statements apply not only to fixtures, but also to test specimens. This is the reasoning behind Columns 1 and 2 of the Criteria Chart, Table I.

EXPERIMENTAL PROOF

Figure 4 shows a cantilevered beam used to model a test specimen. The two blocks and four bolts act as a test fixture, which must always react resonance forces in the specimen. Two conditions of beam stiffness (beam flat and beam on edge) will show how the fixture is affected by changing the length of the test specimen model from 4" to 20". For this case only fixture rigid body rotation f_{RBR} and beam bending f_B frequencies need to be calculated as there are no plate or base twist modes.

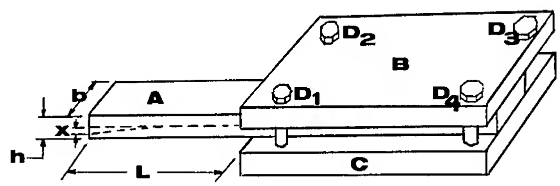


Figure 4 Cantilever beam A represents a test specimen. Blocks B and C, also bolts D_1 through D_4 represent a test fixture. Resonant behavior of the system is predicted (see text) and experimentally verified (see Figure 5).

Fixture rigid body rotation. The rotational moment is TL and stiffness is K_r.

$$f_{RBR} = 3.13\sqrt{\frac{nK_r}{TL}}$$

$$TL = W\left(\frac{L}{2}\right)L_B \quad \text{lb-in}^2\text{/radian}$$

where W = .05L,

L = beam length and

L_B = bolt grip length = 2.5".

The rotational stiffness of the bolts is

$$K_r = \frac{4EI}{L_B}(n) \quad \text{lb/in.}$$

n = 4, in this case, bolts D_1 through D_4 restraining rotation of the beam. I = 0.000975 for a 3/8" bolt, per Table 2, page 10-19, Reference 2. For a 20" beam, f_{RBR} = 271 Hz, while for a 4" beam, f_{RBR} = 1,360 Hz.

Cantilever beam bending.

$$f_B \equiv B_2\sqrt{\frac{386\,E\,I}{WL^4}}$$

where B_2 = 0.56, according to Reference 1, Table 1A.

E = 10 10^6 lb/in^2 for aluminum.

W = 0.05L

I = $bh^3/12$ = 0.0104 inch^4.

For a 20" beam, f_B = 39.7 Hz; for a 4" beam, f_B = 995 Hz. By the same method, the beam response frequencies are calculated when the beam is fastened on its narrow edges. Note that L_B is increased to 3".

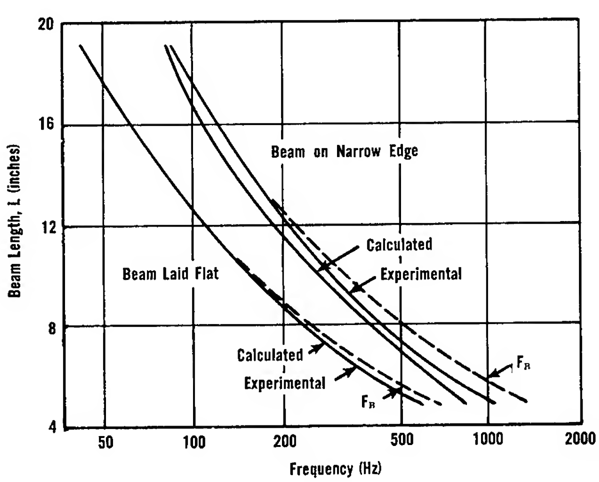


Figure 5 Calculated resonant behavior and experimental results on the system of Figure 4.

Figure 5 graphs the predicted response and experimental results for the two methods of mounting the imaginery specimen. f_{RBR} with the beam flat depended upon mounting bolt stiffness and, combined with bending, compared favorably with frequencies found in the laboratory. Rigid body rotation with the beam on edge coupled with bending as the beam was shortened; this increased the measured response frequency. For beam length 4", f_{RBR} = 1,130 Hz and f_B = 1,990 Hz, yielding f_T approximately 1,000 Hz, as shown in Figure 5. The measured response frequency was 1,470 Hz. One must always consider rigid body rotation. f_T will fall between the calculated f_{RBR} and f_B if the modes are coupled. The mode for beam bending was at the beam root for the 20" beam but moved to the line joining D_3 and D_4 for the 4" beam; this illustrates the change from beam bending to bending plus rotation. Figure 5 also shows how rapidly natural frequency is reduced as fixture size and beam length increase. Realistic design criteria will reflect this change.

CAUSES FOR VARIATION IN MOTION INPUT

Variations in motion intensity relate to fixture and test item resonances. At very low test frequencies, an entire system moves as a single unit; but after the first resonance, varying motions occur. How serious might these be? Figure 6 illustrates test results obtained on a large fixture and large test object (shipboard computer). Three outputs from accelerometers at each mounting point were electrically averaged for control, plotted as the heavy black curve rising to about 55 Hz, then becoming quite constant at 2g. The outputs of those accelerometers experiencing most and least motion at various frequencies are also plotted; the area between these is shaded. Obviously, some mounting points are being overtested and some are being undertested, relative to the average and relative to the test specification. No single control accelerometer location could possibly serve. The motion at each location is the resultant of all forces and of that location's impedance to motion (ability to react the applied forces). These large variations in motion can greatly affect whether an item passes or fails a vibration or shock test.

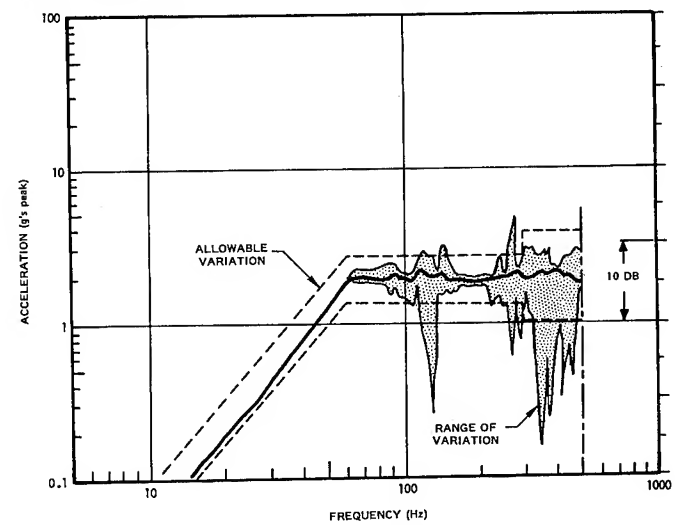


Figure 6 When large objects are tested, some attachment points are overtested and some are undertested. In this example the "spread" exceeds 10:1 at some frequencies. Control accelerometer location greatly affects test outcome.

Testing a large object to frequencies above, say, 500 Hz is usually meaningless for this and other reasons. Measurements on components inside such a computer on board ship would probably show only very low frequencies present, at and below the first major structural resonances of the computer frame.

The problem of locating the control accelerometer faces fixture designers before every test. With high first resonance frequencies, motion

inputs to the several attachment points will be equal over a large frequency range, thus reducing the problem of control variation. This statement is open to any interpretation, of course; the fourth column of the Criteria Chart, Table I, provides quantized guidance.

PROVING THAT DESIGN CRITERIA HAVE BEEN MET

A test item should be mounted on its fixture. A triaxial array of accelerometers should be mounted at each test item attach point. Graph all outputs during a slow sweep through the frequency range, so that all transmissibility peaks, all orthogonal motion and all variations in motion between attach points willbe recorded. (More detailed instructions are found in Section 12 of Reference 2.) If your criteria (Table I or some other source) are not met, find out why. Perhaps a new design is needed, or perhaps minor changes will enable your criteria to be met. Possibly the shaker is inadequate, or auxiliary supports are needed. Or perhaps the test should not be carried to such high frequencies.

When an explanation for a particular resonant peak, for a region of high orthogonal motion or for large variations in motion intensity is sought, remember that the lowest resonance is an f_T summation, according to Dunkerley's Equation, of the natural frequencies of all the beams and plates of the fixture and test article. Somewhat higher in frequency a group of transmissibility peaks and accompanying orthogonal motion and motion variations, caused by the first modes of large individual plates and beams, are often found. Continuing upwards, higher order modes of these, along with first modes of smaller elements are often found; this region is difficult to trace to individual sources and usually cannot be remedied. Fortunately, keeping the first peaks high enough (see the second column of Table I) will usually prevent trouble from these higher modes. This is not to imply that the criteria are easy to meet in all cases; they will often cause difficulty. However, they are generally attainable with good design, particularly when, as at leading laboratories, someone specializes in this design field.

CONCLUSIONS

1. The dynamic behavior of test fixtures is very important to the outcome of vibration and shock tests.
2. Realistic design goals are needed before fixture design commences.
3. If an organization lacks its own design goals, numerical values may be obtained from the Design Criteria Chart found in this paper.
4. Once design and fabrication are complete, but before a test may be commenced, the new fixture's dynamic behavior should be experimentally investigated. If goals are met, fine. If goals are not met, redesign and/or rework may be required. If goals still cannot be met, knowledge of dynamic insuffiences of the fixture often aids in explaining apparent failures during testing.
5. Fixture design is a specialty that should not be entrusted to test technicians or product designers lacking special training.

REFERENCES

1. Karcher, R. F., Shepard, M. L. and Yeager, V; **Standard Environmental Test Methods, Q.E.L. Document QE/C 66-494,** 1 July 1966, Naval Ammunition Depot, Crane, Indiana.
2. Klee, B. J., Kimball, D. V. and Tustin, W; **Vibration and Shock Test Fixture Design.** Tustin Institute of Technology, Santa Barbara, California, 1971.
3. Berkman, H. R.; "Control Point Averaging for Large Specimen Vibration Tests," **Shock and Vibration Bulletin 37, Part 3, January 1968.**

UNIVERSAL OPERATING CHARACTERISTIC CURVES FOR SEQUENTIAL PROBABILITY RATIO TESTS

By David Y. Koo

Astro Communication Laboratory
Gaithersburg, Maryland

ABSTRACT

This paper introduces a simple but practicle technique for constructing an operating characteristic (OC) curve for a sequential probability ratio test (SPRT). Normally in an OC curve for SPRT the probability of acceptance $L(\Theta)$ is a complex function of true equipment MTBF (Θ) and the following parameters: producer's and consumer's risks (α, β) and discrimination ratio (Θ_0/Θ_1 or R). In this paper, the probability of acceptance $L(\Theta)$ is arbitrarily assumed to be an accumulative normal distribution of h variable with scalar and location parameters determined by α, β combination pair. Θ in turn is expressed as a function of h and parameter R, independent of α, β. Consequently, $L(\Theta)$ can be plotted as a straight line on a probability paper for a normal distribution where as the relationship of Θ and h can be graphically represented by a set of nomograph such as that shown on top of Figure 1.

From a reliability engineer's point of view, the essence of the paper is not the simple novel technique of constructing an equivalent OC curve. Rather it is the potential practical applications of such a universal OC curve that deserve our attention. As illustrated in the hypothetical examples 1 and 2 in this paper, the universal OC curve offers us the promise of selecting reliability test plans based on a systematic approach rather than our subjective judgement. Manufacturer can predict the expected risks, maximum risks, expected costs, maximum costs, etc. associated with the chosen range of test plans with greater confidence. Customers can enjoy the benefit of greater varieties of optional test plans with their corresponding price tags roughly defined.

INTRODUCTION

Military Specification MIL-STD-781B Reliability Demonstration Test Plans are designed to demonstrate the MTBF of military and aerospace equipment having exponential distribution of time to failure. There are thirty basic test plans in this specification; ten of these are SPRT plans which are based on the technique for hypothesis testing when neither test time nor number of failure is fixed in advance but is determined during the course of demonstration test. These SPRT plans are most often used because they are usually more efficient than the time or failure terminated tests. To help us to decide which SPRT plan to select, MIL-STD-781B provides an OC curve for each of the SPRT plans; each OC curve shows us the probability of acceptance vs. the true equipment MTBF (Θ). However, quite often we feel that the SPRT plan listed in MIL-STD-781B neither fit our own equipment reliability characteristic nor completely satisfy customer need; consequently, we often wish to propose a modified SPRT plan by redefining producer's risks (α, β), discrimination ratio (R), and specified MTBF (Θ_0). In doing so we need to speculate over a series of OC curves and pick one that is most suited to our need; however, the conventional procedure for plotting an OC curve as shown in the Appendix paragraph 4.2 is too cumbursome to work with. To solve this problem a universal OC curve shown in Figure 1 is introduced.

DERIVATION OF UNIVERSAL OC CURVE FOR SPRT

(For equipment having exponential distribution of time to failure)

As shown in the Appendix, the mathematically derived probability of accepting the equipment when its true MTBF is Θ is:

$$L(\Theta) = \frac{\left(\frac{1-\beta}{\alpha}\right)^h - 1}{\left(\frac{1-\beta}{\alpha}\right)^h - \left(\frac{\beta}{1-\alpha}\right)^h} \tag{A32}$$

$$\Theta = \frac{(R^h - 1)\Theta_0}{h(R-1)} \tag{A34}$$

When α and β values are given we can use equation (A32) to compute a set of points for $L(\Theta)$, h. If we plot these points on a probability paper for a normal distribution, we would get a straight line such as that shown in Figure 1. According to equations (A32) and (A34), when $h=+1$, $L(\Theta)=1-\alpha$ and $\Theta=\Theta_0$; when $h=-1$, $L(\Theta)=\beta$ and $\Theta=\Theta_0/R$. This means as long as we know the values of α, β we can easily plot a straight line OC curve which will readily give us a value of h for each $L(\Theta)$ chosen. Now the remaining task is to substitute the h value into equation (A34) to find the Θ value. Since this is a tedious task, and the reverse process of finding $L(\Theta)$ when Θ is give is even more laborious, a nomograph such as that shown on top of Figure 1 is added to eliminate equation (A34) calculation.

The basic steps for constructing the nomograph is fairly simple. We have already learned that when $h=-1$, $\Theta=\Theta_0/R$ and when $h=+1$, $\Theta=\Theta_0$. When $h=0$ equation (A34) is indeterminant. To solve this problem we can apply L'Hospital rule by differentiating the numberator and denominator of equation (A34) with respect to h separately and take the limit for the ratio as h approches zero; this gives equation (A43) which means $\Theta = m\Theta_0$ when $h=0$. m is the slop of the acceptance or rejection line on a number of failure vs. test time plane for SPRT - see equation (A16). Consequently, the universal reference scale for Θ has three well defined points. The next step is to use equation (A34) to establish a detailed reference scale with a fixed R value; R=10 is chosen because it is quite impracticle to consider a SPRT plan with an R as large as 10. The remaining task is to establish various reference points on the R axis for various R values chosen so that R=10 reference scale Θ values can be projected on the universal reference scale in the proper proportion determined by R. For instance, if R=3, using equation (A17) m is estimated to be 0.548. If we draw a line joining $m\Theta_0$ and $0.548\Theta_0$ on R=10 reference scale we would establish a reference point on R axis. If we want to know what is the value of h and $L(\Theta)$ provided we have chosen MIL-STD-781B

SPRT plan V ($R=3$, $\alpha=\beta=0.1$) and Θ is $0.809\Theta_0$, all we have to do is to connect a straight line between R=3 reference point and $0.809\,\Theta_0$ on R=10 reference scale until it intersects the universal reference scale, and then we draw a verticle line down until it intersects $\alpha=\beta=0.1$ line; see Figure 1. $L(\Theta)$ at this point is 0.8 and h is slightly smaller than 0.667. If we use a cumulative normal distribution table, we would get h (1.282)=0.85 for $L(\Theta)=0.8$; h therefore is about 0.66.

NOTE

As shown in Figure 2 and formulas (A46) through (A50), the universal OC curve can be used for cases where $\alpha \neq \beta$ by simply shifting the h scale to h" scale by an amount equal to c. The α'' and β'' are still equal and $L(\Theta)$ is still equal to 1/2 at h"=0; however, $1-\alpha$ is defined at h=+1 and β is defined at h=-1 which means $\alpha \neq \beta$. Another word we have assumed that:

$$L(\Theta) \doteq \int_{-\infty}^{(h+c)k} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \qquad \text{(A46)}$$

$$= \frac{\left(\frac{1-\beta}{\alpha}\right)^h - 1}{\left(\frac{1-\beta}{\alpha}\right)^h - \left(\frac{\beta}{1-\alpha}\right)^h} \qquad \text{(A32)}$$

Although no formal mathematical proof is given to show that equation (A46) is indeed an equivalent of (A32), Chi-square goodness of fit tests have pointed out that the risk is probably smaller than 1/2 of 1%.

PRACTICAL APPLICATIONS OF UNIVERSAL OC CURVE

Universal OC Curves are not only easy to construct, they are also extremely helpful in test plan selection analysis. To illustrate how easily test plan selection analysis can be carried out, the following hypothetical examples are given:

EXAMPLE 1

Given Conditions :

Θ_0 is specified by customer and the manufacturer knows the actual equipment MTBF (Θ) is roughly within the range $0.4\ \Theta_0$ to Θ_0.

Requirements:

1. Customer Requirements:

   a. R must not exceed 5.
   b. β must not exceed 10%.

2. Manufacturer (MFR) Management Requirement:

   Probability of acceptance must be equal or greater than 55% when Θ is $0.4\Theta_0$.

Problem:

MFR Reliability Group must propose a range of test plans that will satisfy above conditions.

Solution:

For practical reason assume smallest α can be used is 1%.

Step 1: Draw a straight line connecting the reference point R=5 on R scale with $0.4\ \Theta_0$ on R=10 reference scale until it intersects universal reference scale at a point which is roughly $m\ \Theta_0$ in this problem; see Figure 3.

Step 2: Locate a pivotal point which is the intersect of vertical line $m\ \Theta_0$ and horizontal line $L(\Theta)=$ 55%.

Step 3: Draw a straight line connecting pivotal point with the point ($h=1$, $L(\Theta)=1-0.01$). This straight line is L1.

Step 4: Draw a straight line connecting the pivotal point with the point ($h=-1$, $L(\Theta)=10\%$). This is straight line L2.

Step 5: Draw a straight line connecting the points ($h=1$, $L(\Theta)=99\%$), ($h=-1$, $L(\Theta)=10\%$). This is line L3.

Conclusion:

Any OC line completely falls within the shaded region bounded by these three straight lines (L1, L2, L3) will meet all the above requirements, provided R is properly selected. To estimate the price range of these test plans we only need to analyze these three OC lines with R determined by $L(\Theta = 0.4\,\Theta_0) = 55\%$:

L1: R=5, $\alpha=1\%$, $\beta=2\%$
L2: R=5, $\alpha=1-0.94=6\%$, $\beta=10\%$
L3: R=3.7, $\alpha=1\%$, $\beta=10\%$

For instance for a quick rough estimation of test plan involving L3 and R=3.7 we look up Table 2A-1 of HANDBOOK H108, code A9 is selected. Thus, entering Table 2D-1(a) in H108, the following information can be found:

(r_0) - (maximum number of failure when SPRT is truncated)is 27

b_0 - (defined in the Appendix) is 0.8479

b_1 - (defined in the Appendix) is 1.6643

m - (defined in the Appendix) is 0.4843

These figures can be verified using formulas in the appendix.

$E_{\Theta_1}(r)$ - (Expected number of failures when $\Theta=\Theta_1$) is 6.6

$E_s\,(r)$ - (Expected number of failures when $\Theta = m\,\Theta_0$) is 6.0

$E_{\Theta_0}(r)$ - (Expected number of failures when $\Theta=\Theta_0$) is 1.6

From these data expected test time and cost can be estimated; see reference 4 page 208 through 212 for test

time estimation.

EXAMPLE 2

Given Conditions:

Θ_0 is specified by customer and the manufacturer knows from past field failure report or prediction reports the approximate equipment failure rate distribution is normal with mean λ_o and standard deviation(s) equal to $\lambda_o/3$; see Figure 4A. The equivalent Θ distribution is shown in Figure 4B.

Requirements:

1. Customer Requirements:
   a. R must not exceed 5.
   b. β must not exceed 30%.
2. MFR Management Requirements:
   a. The expected probability of passing the reliability test = Expected P(A) $\geq$ 80%
   b. P(A) = 60% at a risk $\leq$10%

Problem:

MFR Reliability Group must propose a range of test plans that will satisfy above conditions.

Solution:

Step 1: Simplification of Mathematical Model for Rough Approximation: For the convenience of analysis let's arbitrarily reduce the model of hypothesis testing to its simplest form - that is if $\Theta \geq \Theta_D$ we call it $\Theta = \Theta_0$ and if $\Theta \leq \Theta_D$ we define it to be Θ_0/R or Θ_1. Thus we can setup the following equations:

See Figure 4 for definitions of equations (EX2-1) and (EX2-2)

By definition $P(A|\Theta=\Theta_0)=1-\alpha$ (EX 2-3)

$P(A|\Theta=\Theta_0/R)=\beta$ (EX 2-4)

Figure 4B $P(\Theta=\Theta_0)=1-2p=P(B)$ (EX 2-5)

$P(\Theta=\Theta_0/R)=2p=P(\bar{B})$ (EX 2-6)

Expected Probability of acceptance

$=\text{Expected } P(A)=P(AB)+P(A\bar{B})$

$=P(A|B)P(B)+P(A|\bar{B})P(\bar{B})$

$=(1-\alpha)(1-2p)+(\beta)(2p)$

$=(1-\alpha)-2p\,[(1-\alpha)-\beta]\geq 80\%$ (EX 2-7)

Step 2: Establish Pivotal Point: To establish pivotal point we can use equation (EX 2-2) and let p'= 10%. From a standard cumulative normal distribution table 3 (F-1)=1.282 or F=1.43. Another word $P(\Theta \leq \frac{1}{1.43}\Theta_0)=P(\Theta \leq 0.7\,\Theta_0) \leq$ 10%. Thus we can establish pivotal point by using $L(\Theta)=60\%$, $\Theta=0.7\Theta_0$, R=5; this will comply with requirement 2b. See Figure 5.

Step 3: Establish R Range: If we repeat the process that is shown in Example 1, we will find that L3 OC line which passes through (h=-1, $L(\Theta)$= 0.3) and the pivotal point in Figure 5 does not comply with (EX 2-7) requirement because the projected α value by L3 line is 0.35 and this means p must be negative - which is impossible in reality. To bypass this pitfall let us arbitrarily choose the range of R between 1.5 to 5. If we let R=1.5 in equation (EX 2-1), p would be equal to 0.0668. Substitute this p value into (EX 2-7):

$.2-.8664\alpha+.136\beta\geq 0$ (EX 2-8)

Since: Since β is $\leq$ 0.3, 0.136β is much smaller than .2 so if we ignore the third term in (EX 2-8); we can establish a maximum limit for α.

$.2\geq .8664\alpha$ $\alpha\leq 0.231$ or $\alpha_{max}=0.231$

Step 4 Establish 3 OC Lines (that will satisfy all requirements).

L1: Connect a straight line between (h=1, $L(\Theta)$=0.99) and pivotal point.

L2: Connect a straight line between (h=-1,$L(\Theta)$=0.3) and (h=+1, $L(\Theta)$=0.99)

L3': Connect a straight line between (h=1, $L(\Theta)=1-\alpha_{max}$=.769) and pivotal point.

L4: Connect a straight line between (h=-1, $L(\Theta)$=0.01) and pivotal point.

Any OC line falls completely within the shaded region will comply with all the requirements, provided R must be within 1.5 to 5 and chosen properly.

REFERENCES

1. Mood, A.M. and Graybill, F.A., "Introduction to the Theory of Statistics, "New York, McGraw Hill Book Co., 1963
2. Wald, A., "Sequential Analysis, "New York, John Wiley & Sons, 1947
3. "Reliability Testing, Exponential Distribution, "MIL-STD-781B, November 15, 1967
4. Myers, R., "Reliability Engineering for Electronic Systems, "New York, Wiley & Sons, 1964
5. "Sampling Procedures and Tables for Life and Reliability Testing, Based on Exponential Distribution "HANDBOOK H108, April 29, 1960

APPENDIX

1. Introduction

In order to understand the universal OC curve and its derivation procedure better, it is necessary to retrace the process from which the SPRT was derived. Consequently, an extensive mathematical derivation of SPRT plan (based on exponential life distribution) is presented in the appendix to serve the dual purpose of ready reference for this paper as well as a convenient digest for the SPRT plans listed in MIL-STD-781B and HANDBOOK H108.

2. Definitions

Θ_0 - Specified MTBF
Θ_1 - Minimum Acceptable MTBF
α - Producer's Risk = P (REJ H_0 | Θ_0)
β - Consumer's Risk=P (Accept H_0 | Θ_1)
R - Discrimination Ratio = Θ_0 / Θ_1

3. Hypothesis Testing Based on Exponential Distribution

T1	T2	T3	......	Tn	
	1st	2nd	3rd	(n-1)th	n th Failure

T_1, T_2 .... T_n are accumulative equipment test times between failures.

Joint distributions of T's = $f(T_1, T_2, \ldots\ldots, T_n)$

$$= \prod_{i=1}^{n} f(T_i;\Theta) = \frac{1}{\Theta^n} \text{EXP}\left[-\frac{1}{\Theta}\sum_{i=1}^{n} T_i\right] \quad \text{(A1)}$$

Equation (A1) is true because T's are statistically independent.

H_0 (Null Hypothesis): $\Theta=\Theta_0$
H_1 (Alternative Hypothesis): $\Theta=\Theta_1$

$$B < \frac{\prod_{i=1}^{n} f(T_i;\Theta_1)}{\prod_{i=1}^{n} f(T_i;\Theta_0)} < A \quad \text{(A2)}$$

Note that when this ratio is close to 1 test continues. If the ratio is >> 1, H_0 will be rejected. If ratio <<1, H_0 will be accepted.

$$B < \frac{\left[\frac{R}{\Theta_0}\right]^n \text{EXP}\left[-\frac{R}{\Theta_0}\sum_{i=1}^{n} Ti\right]}{\left[\frac{1}{\Theta_0}\right]^n \text{EXP}\left[-\frac{1}{\Theta_0}\sum_{i=1}^{n} Ti\right]} < A \quad \text{(A3)}$$

$$B < R^n \text{EXP}\left[-\frac{(R-1)}{\Theta_0}\sum_{i=1}^{n} T_i\right] < A \quad \text{(A4)}$$

Let $B < L_n < A$ (A5)

$$L_1 = \frac{f(T_1;\Theta_1)}{f(T_1;\Theta_0)} \quad \text{(A6)}$$

$$L_n = \frac{\prod_{i=1}^{n} f(T_i;\Theta_1)}{\prod_{i=1}^{n} f(T_i;\Theta_0)} \quad \text{(A7)}$$

$$= P(L_1 \geq A) + P(B<L_1<A,\ L_2 \geq A) + P(B<L_1<A,\ B<L_2<A, L_3 \geq A) + \ldots\ldots \quad \text{(A8)}$$

$$= P(L_1 \leq B) + P(B<L_1<A,\ L_2 \leq B) + P(B<L_1<A,\ B<L_2<A,\ L_3 \leq B) + \ldots\ldots \quad \text{(A9)}$$

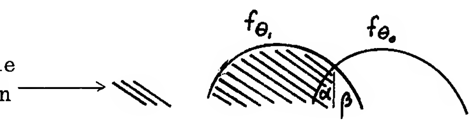
The actual determination of A & B from above equations is too elaborate; however, the following approximate solution can be found by rationalization. Suppose that L_n is a continuous function of a continuous variate n so that at some value of n L_n first equals A or B:

Condition 1

Assume true equipment MTBF $=\Theta_1$ so that after n failures L_n is just =A so H_0 is rejected.

$$\frac{P(\text{Test sample in } f_{\Theta_1})}{P(\text{Test sample in } f\Theta_0)} = L_n = A \doteq \frac{1-\beta}{\alpha} \quad \text{(A10)}$$

Test Sample Distribution

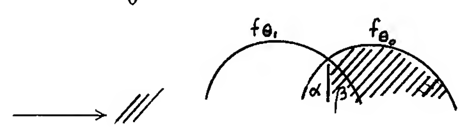


Condition 2

Assume true equipment MTBF$=O_0$ so that after n failures L_n is just = B so H_0 is accepted.

$$\frac{\text{Probability test sample in } f\Theta_1}{\text{Probability test sample in } f\Theta_0} = L_n = B \doteq \frac{\beta}{1-\alpha} \quad \text{(A11)}$$

Test Sample Distribution



Substitude $A \doteq \frac{1-\beta}{\alpha}$, $B \doteq \frac{\beta}{1-\alpha}$ into equation (A4)

$$\frac{\beta}{1-\alpha} < R^n \text{EXP}\left[-\frac{(R-1)}{\Theta_0}\sum_{i=1}^{n} T_i\right] < \left(\frac{1-\beta}{\alpha}\right) \quad \text{(A12)}$$

$$\left(\frac{\beta}{1-\alpha}\right)(R^{-n}) < \text{EXP}\left[-\frac{(R-1)}{\Theta_0}\sum_{i=1}^{n} Ti\right] < \left(\frac{1-\beta}{\alpha}\right)R^{-n} \quad \text{(A13)}$$

$$\ln\left(\frac{\beta}{1-\alpha}\right)+\ln(R^{-n})< -\frac{(R-1)}{\Theta_0}\sum_{i=1}^{n} T_i<\ln\left(\frac{1-\beta}{\alpha}\right)+\ln(R^{-n}) \qquad (A14)$$

Note: $(\ln x=2.3026 \log_{10} x)$

$$n\left(\frac{2.3}{R-1}\right)\log_{10} R+ \frac{2.3}{R-1}\log_{10}\left(\frac{1-\alpha}{\beta}\right)> \frac{\sum_{i=1}^{n} T_i}{\Theta_0} >$$

$$n\left(\frac{2.3}{R-1}\right)\log_{10} R-\frac{2.3}{R-1}\log_{10}\left(\frac{1-\beta}{\alpha}\right) \qquad (A15)$$

$$mx+b_0>y>mx-b_1 \qquad (A16)$$

x=n=Total number of failures
y=Total test time in multiples of Θ_0
$R=\Theta_0/\Theta_1$

When $y > mx+b_0$ Accept H_0
When $y < mx-b_1$ Reject H_0

$$m=\left(\frac{2.3}{R-1}\right)\log_{10} R \qquad (A17)$$

$$b_0=\frac{2.3}{R-1}\log_{10}\left(\frac{1-\alpha}{\beta}\right) \qquad (A18)$$

$$b_1=\frac{2.3}{R-1}\log_{10}\left(\frac{1-\beta}{\alpha}\right) \qquad (A19)$$

4. Study of OC Curve

4.1 OC Curve is used to estimate the probability of accepting null hypothesis (H_0) when the true equipment MTBF mean is specified.

Let the true equipment MTBF mean=Θ, thus $f(t;\Theta)=\frac{1}{\Theta}e^{-t/\Theta}$

$$\text{Let}\quad \frac{g(t;\Theta)}{f(t;\Theta)}=\left[\frac{f(t;\Theta_1)}{f(t;\Theta_0)}\right]^h \qquad (A20)$$

$0\le t\le\infty$
t - random variable test time

$$\text{and let } \phi(u)=\int_{-\infty}^{\infty}\left[\frac{f(t;\Theta_1)}{f(t;\Theta_0)}\right]^u f(t;\Theta)dt \qquad (A21)$$

Let $g(t;\Theta)$ be a density function - that is:

$\phi(u=h)=1$

Thus using equations (A1) and (A21):

$$\phi(u)=\int_0^{\infty}\left(\frac{\Theta_0}{\Theta_1}\right)^u\left[\frac{e^{-t/\Theta_1}}{e^{-t/\Theta_0}}\right]^u\left(\frac{1}{\Theta}\right)e^{-t/\Theta}dt \qquad (A22)$$

Note: t- test time can not be negative!

$$\phi)u=h)=1=\left(\frac{\Theta_0}{\Theta_1}\right)^h\frac{1}{\Theta}\int_0^{\infty}e^{-\left[\frac{\Theta_1\Theta_0+h\Theta\Theta_0-\Theta\Theta_1 h}{\Theta_1\Theta_0\Theta}\right]}dt \qquad (A23)$$

Substitute $R=\Theta_0/\Theta_1$ into (A23)

$$R^h-(R-1)\frac{h\Theta}{\Theta_0}-1=0 \qquad (A24)$$

$$\frac{1-R^h}{h(1-R)}=\frac{\Theta}{\Theta_0} \qquad (A25)$$

Now we can setup an equivalent hypothesis test:

We can rewrite equation (A2) or (A5) as:

$$B^h<\left[L_n\right]^h<A^h \qquad (A26)$$

$$B^h<(L_n)^h\frac{f(t_1;\Theta)\,f(t_2;\Theta)\ldots\ldots f(t_n;\Theta)}{f(t_1;\Theta)\,f(t_2;\Theta)\ldots\ldots f(t_n;\Theta)}<A^h \qquad (A27)$$

This is equivalent new hypothesis testing:

H'_0: True density function is $f(t;\Theta)$

H'_1: True density function is $g(t;\Theta)$

$$B^h<\frac{g(t_1;\Theta)g(t_2;\Theta)\ldots\ldots\ldots g(t_n;\Theta)}{f(t_1;\Theta)f(t_2;\Theta)\ldots\ldots\ldots f(t_n;\Theta)}<A^h \qquad (A28)$$

Note that equation (A28) is equivalent to equation (A2) except that we have added Θ (the true equipment MTBF mean) and h (a function of Θ in the given test plan defined by Θ_0, R); see equation (A24) or (A25). A & B are functions of α and β defined in equations (A10) and (A11).

Now again use the same logic stated in conditions 1 & 2 in paragraph 3.

$$A^h=\frac{1-\beta'}{\alpha'} \qquad (A29)$$

$$B^h=\frac{\beta'}{1-\alpha'} \qquad (A30)$$

α'=probability rej. H'_0 when $f(t;\Theta)$ is true
β'=probability accept H'_0 when $g(t;\Theta)$ is true
From equations (A29) and (A30):

$$\alpha'=\frac{1-B^h}{A^h-B^h} \qquad (A31)$$
= Probability of rej. H'_0 or H_0 when the true equipment distribution is $f(t;\Theta)$

4.2 From equations (A25) and (A31) we can setup a series of equations that can help us to construct an OC curve for any given test plan.

$L(\Theta)$ = Probability of accepting H'_0 or H_0 when true equipment distribution is $f(t;\Theta)$ is:

$$1-\alpha' = 1 - \frac{1-B^h}{A^h - B^h} = \frac{A^h - 1}{A^h - B^h} \quad \text{(A32)}$$

When $\alpha=\beta$, $A^h = 1/B^h$, thus (A32) becomes:

$$L(\Theta \mid \alpha=\beta) = \frac{1}{1+B^h} \quad \text{(A33)}$$

From (A25) $\quad \Theta = \dfrac{(R^h-1)\Theta_0}{h(R-1)} \quad$ (A34)

From (A34) when $\dfrac{\Theta}{\Theta_0} = 1$, $h=1$

Substitute into (A32):

$$L(\Theta \mid \Theta=\Theta_0) = 1-\alpha \quad \text{(A35)}$$

Form (A34) when $\dfrac{\Theta}{\Theta_0} = \dfrac{1}{R}$, $h=-1$

Substitute into (A32):

$$L(\Theta \mid \Theta=\Theta_1) = \beta \quad \text{(A36)}$$

Let $h=-\infty$ & solve (A32) & (A34):

$$L(\Theta \mid \Theta=0) = 0 \quad \text{(A37)}$$

Let $h=\infty$ & solve (A32) & (A34):

$$L(\Theta \mid \Theta=\infty) = 100\% \quad \text{(A38)}$$

$$L(\Theta \mid \Theta=m\Theta_0) = \frac{b_1}{b_1+b_0} \quad \text{(A39)}$$

$$L(\Theta \mid \Theta=m\Theta_0,\ \alpha=\beta) = 50\% \quad \text{(A40)}$$

b_1, b_0 are defined in equations (A18) & (A19)

Proof: Let $h=0$ we can not use (A32) directly because:

$L(\Theta \mid h=0) = \dfrac{A^0-1}{A^0-B^0}$ is indeterminant

If we apply L'Hospital rule by differentiating numerator and denominator separately with respect to h and take the limit:

$$L(\Theta \mid h=0) = \lim_{h\to 0}\left[\frac{A^h \log_e A}{A^h \log_e A - B^h \log_e B}\right] = \frac{\ln A}{\ln A - \ln B} \quad \text{(A41)}$$

Substitute equations (A10), (A11), (A18), (A19) into (A41)

$$L(\Theta \mid h=0) = \frac{2.30\log_{10} A}{2.30\log_{10} A - 2.30\log_{10} B} = \frac{b_1(R-1)}{b_1(R-1)+b_0(R-1)}$$

$$= \frac{b_1}{b_1+b_0} \quad \text{(A42)}$$

Note: $-\ln B \doteq -\ln \dfrac{\beta}{1-\alpha} = +\ln \dfrac{1-\alpha}{\beta}$

$b_0(R-1) = \ln\left(\dfrac{1-\alpha}{\beta}\right)$

When we apply L'Hospital rule to (A34) and take the limit as h approaches zero:

$$\Theta = \lim_{h\to 0}\left[\frac{\Theta_0 R^h \ln R}{R-1}\right] = \Theta_0\left(\frac{\ln R}{R-1}\right) = \Theta_0 m \quad \text{(A43)}$$

Definition of m is defined in (A17)

4.3 Example of OC Curve Plotting (using equations in Section 4.2)

(Given) Test Plan VI in MIL-STD-781B
$\alpha, \beta = 10\%$, $R=5$

(Find) OC Curve

(Solution) $B = \dfrac{\beta}{1-\alpha}$ (A11)

$$B = \frac{0.1}{1-0.1} = \frac{1}{9}$$

Let $h=1$. This means $\Theta=\Theta_0$ according equation (A34)

$L(\Theta\ \Theta=\Theta_0) = 1-\alpha \quad$ (A35) $= 1-0.1 = 0.9$

Let $h=-1$. This means $\Theta=\Theta_1$ or $\Theta/\Theta_0 = \dfrac{1}{R}$ according (A34)

Let $(\Theta\ \Theta=\Theta_1) = \beta = 0.1$ (A36)

$L(\Theta\ \Theta=0) = 0$ (A37)

$L(\Theta\ \Theta=\infty) = 100\%$ (A38)

$L(\Theta\ \Theta=m\Theta_0,\ \alpha=\beta) = 50\%$ (A40)

$m = \left(\dfrac{2.3}{R-1}\right)\log_{10} R \quad$ (A17) $= \dfrac{2.3}{5-1}\log_{10}(5) = \dfrac{2.3}{4}(.7) -$

0.4025

$\Theta = 0.4025\Theta_0$

From the above five points we can construct an OC curve that is very similar to that in P68 of MIL-STD-781B. The difference is due to α, β of test plan VI is not exactly 10% as stated in page 60 of MIL-STD-781B.

4.4 Universal SPRT OC Curve (when α equal to β). If we assume (A44) is true, $L(\Theta)$ vs h plot on an accumulative normal distribution paper will appear as a straight line; see Figure 1.

$$L(\Theta) \doteq \int_{-\infty}^{hk} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz \quad \text{(A44)}$$

Note Θ and h relation is defined in (A33) and (A34).

K is a function of α, β shown in equation (A45)

K	1.645	1.282	1.0365	0.8418	0.5244
$\alpha = \beta$	0.05	0.10	0.15	0.20	0.30

$$1-\alpha \doteq \int_{-\infty}^{k} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \qquad \text{(A45)}$$

Note that equation (A44) is an equivalent of (A33). Although no formal proof is given in this report. A Chi-square goodness of fit test shown below has demonstrated that the risk of assuming equation (A44) is equation (A33) is much less than 1/2 of 1%.

1	h	-1.282	-1.0	-0.5	-0.334	0	+0.334	+0.5	+1.0	+1.282
2	A	0.05	0.10	0.25	0.324	0.5	.666	0.75	0.90	0.95
3	B	0.055	0.10	0.261	0.335	0.5	.676	0.739	0.90	0.945
4	$\frac{(A-B)^2}{A}$	$\frac{(.005)^2}{.05}$	0	$\frac{(.011)^2}{.25}$	$\frac{(.001)^2}{.324}$	0	$\frac{(.01)^2}{.666}$	$\frac{(.011)^2}{.75}$	0	$\frac{(.005)^2}{.95}$

Sum of row 4 << $\chi^2_{.005}$ (for 8 d.f.)

4.5 Universal SPRT OC Curve (when α not necessarily equal to β).

$$L(\Theta) = \int_{-\infty}^{(h+c)k} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \qquad \text{(A46)}$$

Note: For α, β <1/2 $\qquad$ $Z_{1-\alpha}$ always positive
-1<C<1; K$\geq$0 $\qquad$ Z_β always negative

$$1-\alpha \doteq \int_{-\infty}^{(1+c)k} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \int_{-\infty}^{z_{1-\alpha}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \qquad \text{(A47)}$$

$$\beta = \int_{-\infty}^{(-1+c)k} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \int_{-\infty}^{z_\beta} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz \qquad \text{(A48)}$$

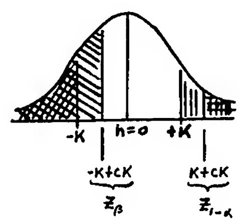
Derivations of equations (A46):

As long as equations (A44) and (A45) are true, it is obvious that we can select $\alpha \neq \beta$ by simply shifting the origin on the h scale. As shown in the Figure 2, the straight line L1 represent an OC curve that $\alpha = \beta$; conquently, equation (A44) and (A45) can be used. This is same as using equation (A46) through (A48) when C=0. We can then arbitrarily define α'' and β'' in the following manner:

$$\alpha'' = \int_{k}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \int_{-\infty}^{-k} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \beta''$$

$$C = \frac{Z_{1-\alpha} + Z_\beta}{Z_{1-\alpha} - Z_\beta} \qquad \text{(A49)}$$

$$K = \frac{Z_{1-\alpha} - Z_\beta}{2} \qquad \text{(A50)}$$



(A49) and (A50) - derived by using equations (A47) and (A48) and solve:

$(1+C)K = Z_{1-\alpha}$

$(-1+C)K = Z_\beta$

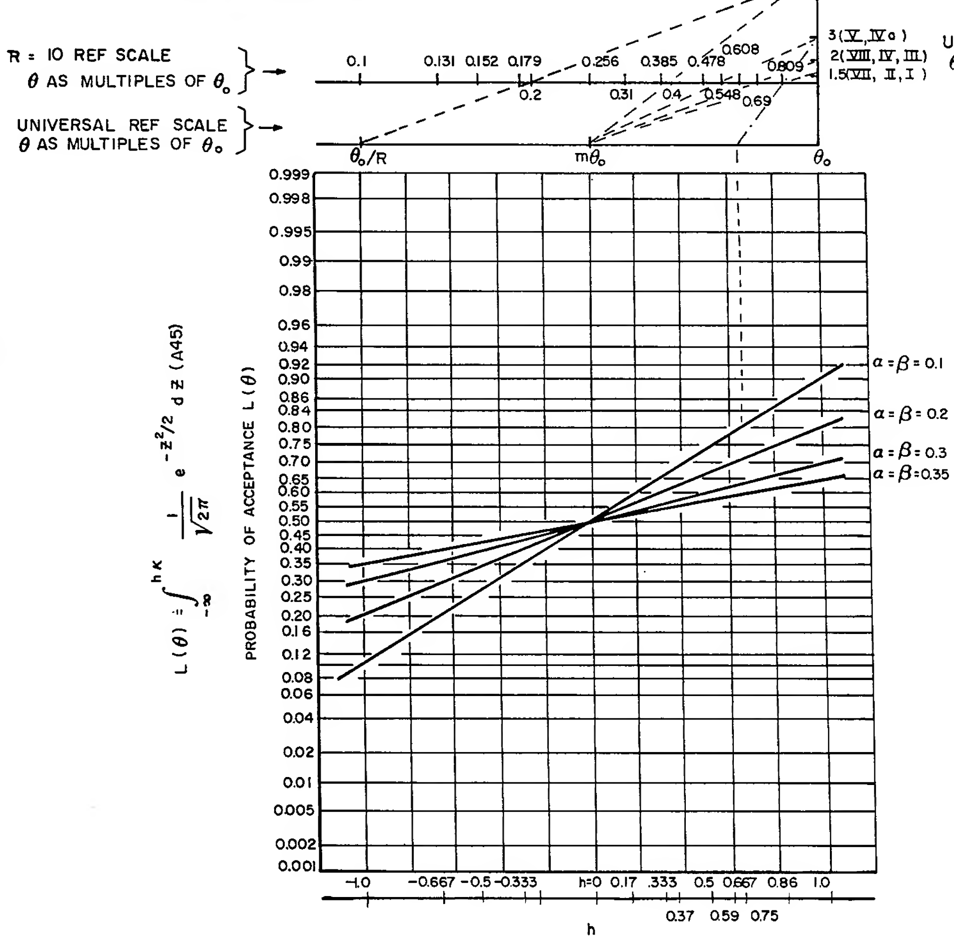
$/\!/\!/ \leftarrow \beta''$ $\quad \backslash\backslash\backslash \leftarrow \beta$ $\quad ||| \leftarrow \alpha''$ $\quad \equiv \leftarrow \alpha$ $\quad \alpha'' = \beta''$ $\quad \alpha \neq \beta$

As shown in Figure 2 if we have shifted L1 by an amount equal to C, we get L2. In this case $\alpha \neq \beta$ but α'' still equal to β''. This means L(Θ) vs. h plot on accumulative normal probability paper will always appear as a straight line regardless what α, β values are.

MIL-STD-781B SPRT PLANS

	I	II	III	IV	IVa	V	VI	VII	VIII	IX
α,β	0.1	0.2	0.1	0.2	0.2	0.1	0.1	0.3	0.3	0.35,0.4
(θ_0/θ_1) OR R	1.5	1.5	2	2	3	3	5	1.5	2	1.25
m	0.809	0.809	0.69	0.69	0.548	0.548	0.4	0.809	0.69	0.902

α,β	0.1	0.2	0.3	0.35
κ	1.282	0.8418	0.5244	0.385



$$L(\theta) \doteq \int_{-\infty}^{h\kappa} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz \quad (A45)$$

FIGURE 1. UNIVERSAL O.C. CURVE FOR SPRT ($\alpha = \beta$).

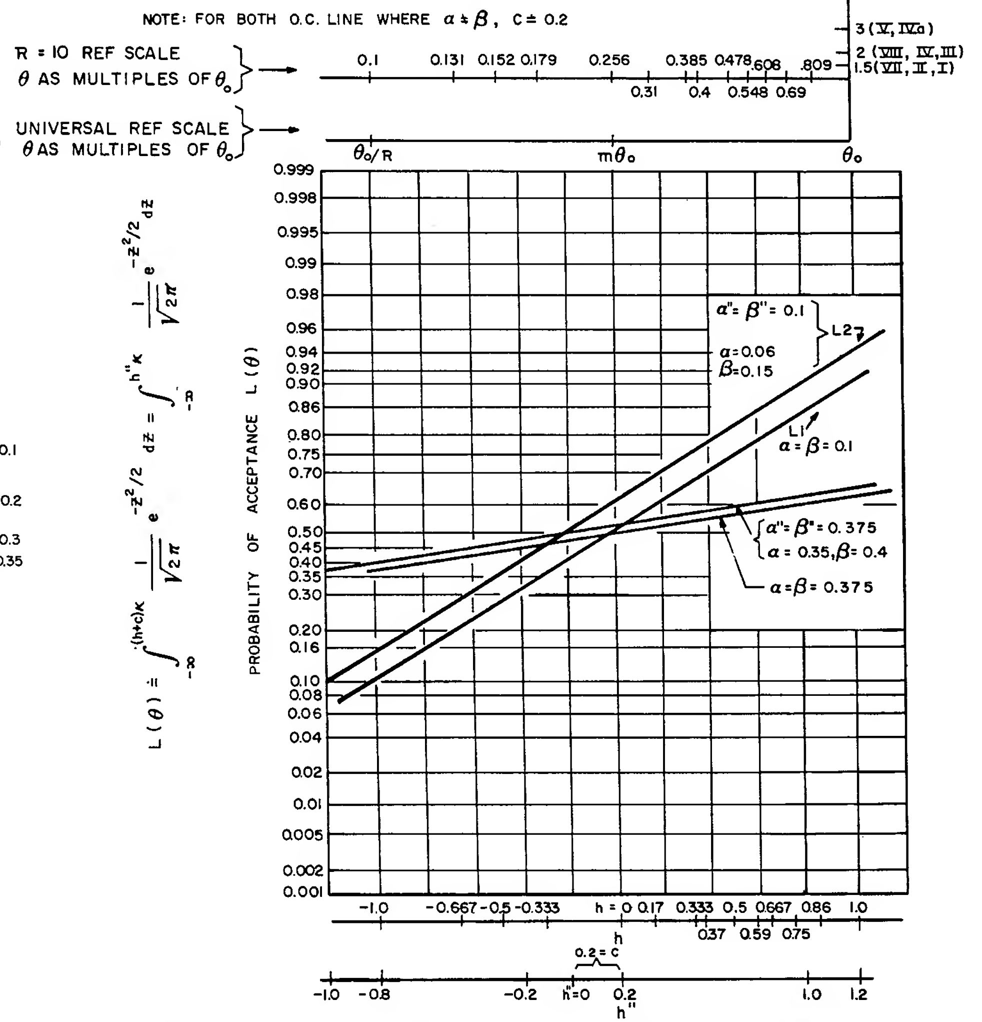
$$L(\theta/h = +1) \doteq 1-\alpha \doteq \int_{-\infty}^{z_{1-\alpha}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz \quad (A47) \qquad C = \frac{z_{1-\alpha}+z_\beta}{z_{1-\alpha}-z_\beta} \quad (A49)$$

$$L(\theta/h = -1) \doteq \beta \doteq \int_{-\infty}^{z_\beta} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz \quad (A48) \qquad \kappa = \frac{z_{1-\alpha}-z_\beta}{2} \quad (A50)$$

$L(\theta/h'' = +1) \doteq 1 - \alpha''$

$L(\theta/h'' = -1) \doteq \beta'' = \alpha''$ ALWAYS BECAUSE NORMAL DISTRIBUTION IS SYMETRICAL ABOUT ITS MEAN.

NOTE: FOR BOTH O.C. LINE WHERE $\alpha \neq \beta$, $C \doteq 0.2$



$$L(\theta) \doteq \int_{-\infty}^{(h+c)\kappa} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz = \int_{-\infty}^{h''\kappa} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$$

FIGURE 2. GENERAL UNIVERSAL O.C. CURVE FOR SPRT (α NOT NECESSARILY EQUAL TO β).

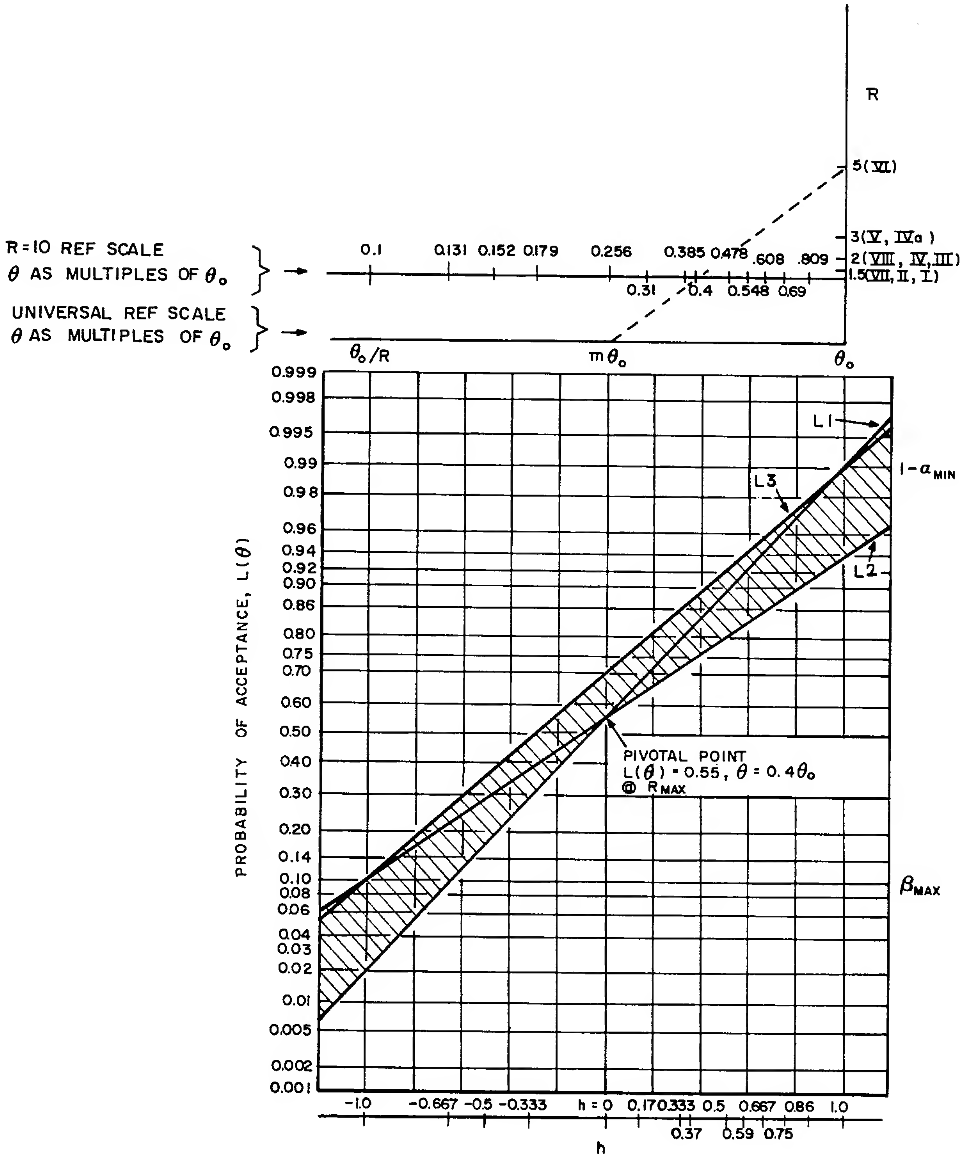


FIGURE 3. RANGE OF UNIVERSAL O.C. CURVES THAT COMPLY WITH EXAMPLE I REQUIREMENTS.

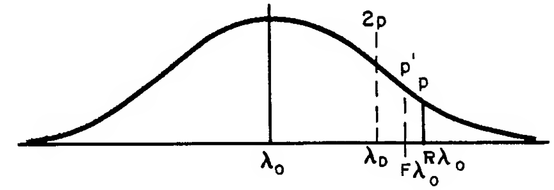


FIGURE 4A. HYPOTHETICAL ESTIMATED EQUIPMENT FAILURE RATE (λ) DISTRIBUTION FOR EXAMPLE 2.

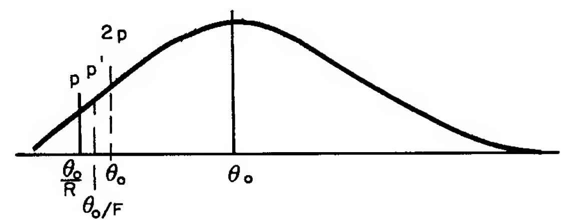


FIGURE 4B. HYPOTHETICAL ESTIMATED EQUIPMENT MTBF (θ) DISTRIBUTION FOP EXAMPLE 2.

$$P \doteq \int_{\frac{\lambda_o (R-1)}{s}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \doteq \int_{3(R-1)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \qquad \text{(EX 2-1)}$$

$$P' \doteq \int_{3(F-1)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \qquad \text{(EX 2-2)}$$

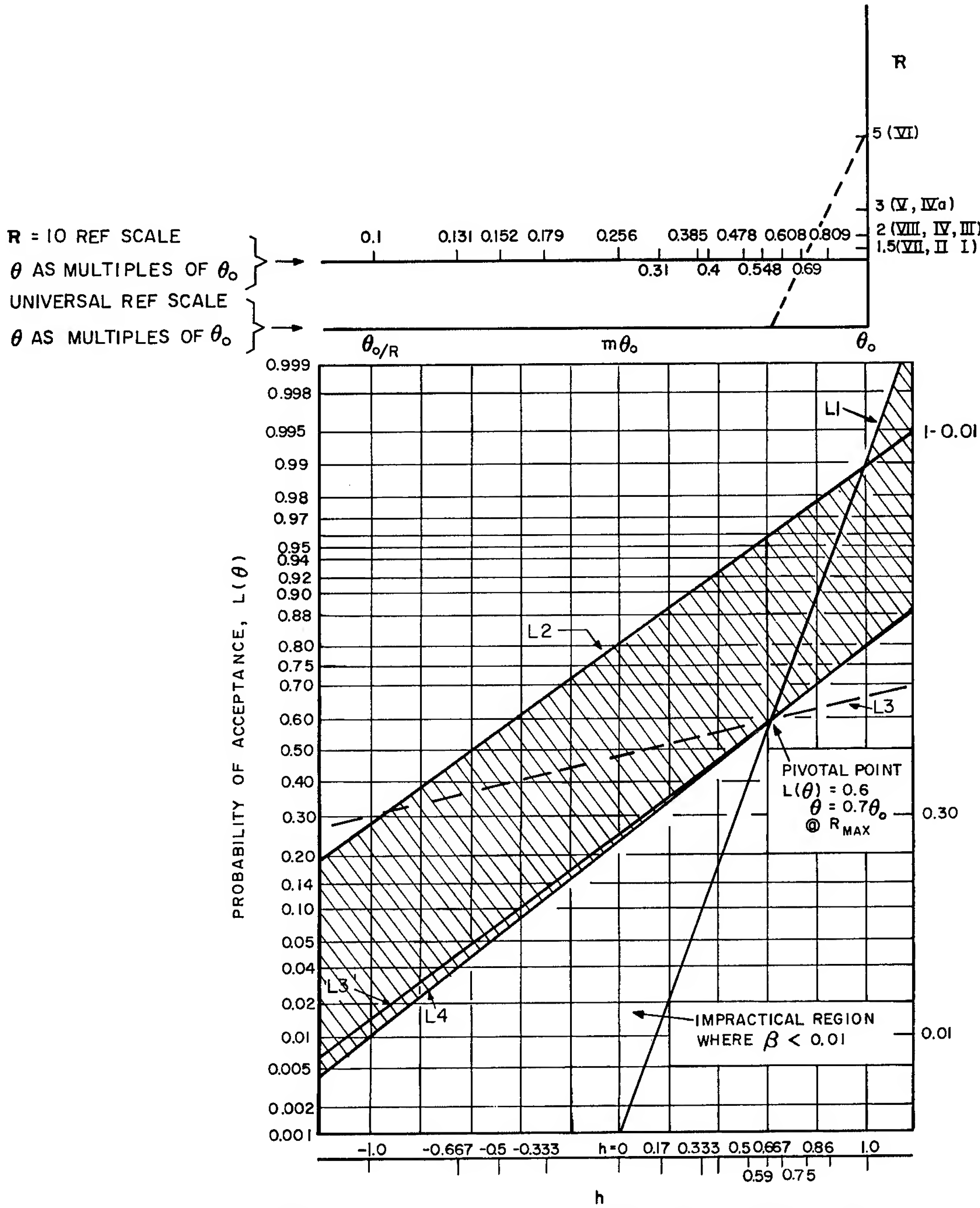


FIGURE 5. RANGE OF UNIVERSAL O.C. CURVES THAT COMPLY WITH EXAMPLE 2 REQUIREMENTS.

ACCELERATED TESTING OF AIR-TO-AIR GUIDED MISSILES

INDEX SERIAL NUMBER – 1054

T. W. Elliott
Systems Test Division, Naval Missile Center
Point Mugu, Calif.

Abstract

Most environmental testing of air-to-air guided missiles, particularly vibration testing, is based upon an accelerated model whereby the stresses to be encountered by the missile over a long period of time in service are simulated in a shorter time in a laboratory test. All too often, when designing the test, the acceleration model is either not stated or only tacitly assumed and only rarely is the model investigated for accuracy. The purpose of this paper is to review one model proposed in the literature and show the agreement obtained from a test program performed at the NAVMISCEN (Naval Missile Center) on a current Fleet air-to-air guided missile.

The NAVMISCEN recently completed a test program on a current Fleet air-to-air guided missile where the objective was to demonstrate that in-service captive flight reliability could be adequately predicted from testing in a laboratory at higher than service mean levels. Four missiles were tested at two different vibration levels above in-service conditions. For each of these levels an MTBF was calculated from the results and the acceleration model determined. From this model, reliability at in-service conditions was predicted. The predicted reliability compared favorably with indicated reliability from Fleet data and the acceleration model compared favorably with a model suggested by A. J. Curtis of Hughes Aircraft Company. Results from a third and higher vibration level demonstrated a definite upper limit for accelerated testing.

From the results of this test program it was concluded that high confidence can be placed on results of accelerated testing when the model has been properly investigated. Considering the current increased emphasis on laboratory testing and reliability determination it is recommended that investigation of the acceleration model for each system be undertaken as early as possible.

Introduction

Most environmental testing of air-to-air guided missiles, particularly vibration testing, is based upon an acceleration model whereby the stresses to be encountered by the missile over a long period of time in service are simulated in a shorter period of time in a laboratory test. The objectives of these tests usually fall in one of the two categories of qualification (specification compliance) or reliability (mean time between failures) testing where real-time and real-level testing of a missile, which may be captive-carried for hundreds of hours, becomes overly expensive. All too often, however, when the test is initially designed this acceleration model is either not stated or only inherently assumed, and only rarely is the model investigated for accuracy. For example, a missile may be designed to withstand 500 hours of nominal captive flight and, presumably, to demonstrate this capability, the missile is vibrated at some higher than nominal level for 2 hours. Thus a relationship between 500 hours at nominal level and 2 hours at the higher level is implied but rarely stated and even more rarely ever investigated for accuracy. Without adequate knowledge of this relationship, as it applies to a particular system, serious errors can be made which will become evident only after years of in-service operation. The purpose of this paper is to review one model proposed in the literature and show the agreement obtained from a test program performed at the NAVMISCEN (Naval Missile Center) on a current Fleet air-to-air guided missile.

Background

In Shock and Vibration Monograph No. 8, "Selection and Performance of Vibration Tests,"[1] A. J. Curtis, N. Tinling, and H. Abstein present an excellent discussion of accelerated vibration testing in which is presented the following equation for exaggeration factor in random vibration testing (p. 91)

$$\left(\frac{W_1}{W_2}\right)^{b/n} = \frac{T_2}{T_1}$$

where W = Power spectral density level
b = Measure of slope of S-N curve
n = Damping-stress exponent
T = Time
1, 2 = Test condition subscripts

It is further stated that the value of b ranges from 3 to 25, with a representative value of 9, and the value of n is 2.4 for stresses below 80 percent of the endurance limit and 8 for stresses above 80 percent of the endurance limit.

When a mathematical model, such as the above, is used to define an accelerated test, the decision must first be made as to what values to assign the constants b and n. For an air-to-air guided missile system, there usually exists a specification on captive-flight life such as "the missile must withstand 500 hours of captive flight." Thus it is reasonable to assume that a designer will design his product such that the captive-flight-induced forces will induce stresses that are less than the endurance limit of the product to avoid underdesign but only slightly less than the endurance limit to avoid overdesign. Therefore, in an accelerated test, it must be assumed that the higher than nominal input forces of the accelerated levels must be above 80 percent of the endurance limit, and the proper value for the constant n is 8 for this case.

Referring back to the above mathematical model, it is convenient, for the purpose of this paper, to rearrange it slightly to reflect acceleration in terms of g level instead of spectral density level. Since spectral density is proportional to the square of the g level, or

$$W \approx g^2$$

then

$$\left(\frac{W_1}{W_2}\right)^{b/n} = \left(\frac{g_1}{g_2}\right)^{2\ b/n} = \frac{T_2}{T_1}$$

Substituting b = 9, n = 8,

$$\frac{T_2}{T_1} = \left(\frac{g_1}{g_2}\right)^{(2 \times 9)/8} = \left(\frac{g_1}{g_2}\right)^{2.25}$$

Thus the ratio of times at two levels is equal to the inverse ratio of the levels raised to the 2.25 power.

The next part of this paper will review a test program performed by the NAVMISCEN in which, from the failure data, a model relating mean time between failures (MTBF) to vibration level was derived with good agreement to the above model.

Accelerated Failure Rate Test Program

Introduction

As part of an intensive effort to improve effectiveness of air-to-air missiles, NAVAIRSYSCOM (Naval Air Systems Command) sponsored a reliability improvement program in which the NAVMISCEN conducted a series of environmental tests on a sample of missiles. The objective of this effort was to measure the MTBF, in a laboratory program, of a sample of missiles from the current inventory as a baseline for determining reliability improvement in later improved designs. Environmental simulation test techniques (accelerated failure rate tests) were used to pursue this objective. The test criteria were:

1. Missile failures that occurred during the accelerated failure rate test must be similar to failures that occurred during service.

2. The MTBF of the accelerated failure rate tests need not be identical to that of the missiles in service because the objective of the program was comparison before and after improvement; however, adjustment of the MTBF value, through mathematical modeling, to a value corresponding to in-service values was desirable.

Method. Four missiles from the current Navy inventory were randomly selected, instrumented, and subjected to the accelerated failure rate tests. The experimental design used for the accelerated failure rate tests is shown in table 1.

Table 1. Test Matrix

Test	2g Rms Vibration 180 Minutes	4g Rms Vibration 45 Minutes	Test	2g Rms Vibration 180 Minutes	4g Rms Vibration 45 Minutes
Longitudinal Axis					
1		C	25	C	
2	A		26		A
3		B	27	B	
4	D		28		D
5	C		29		C
6		D	30	D	
7	B		31		B
8		A	32	A	
Wingplane 1-3					
9		B	33	B	
10	D		34		D
11		C	35	C	
12	A		36		A
13	B		37		B
14		A	38	A	
15		C	39	C	
16	D		40		D
Wingplane 2-4					
17	B		41		C
18		D	42	A	
19	C		43		B
20		A	44	D	
21		C	45	C	
22	D		46		A
23		B	47	B	
24	A		48		D

Vibration levels were based on vibration data obtained from captive-flight measurements of samples of these missiles flown at the NAVMISCEN. The spectrum chosen for level 1 was a smoothed version of the highest of the measured acceleration spectral densities. Level II was twice the rms g value of level 1 or four times the spectral density. Figure 1 shows the vibration spectra used for the accelerated failure rate tests.

Discussion. During the progress of the test program it was necessary to deviate from the original test plan for a number of reasons. The reasons for these deviations were:

1. At the beginning of the test program the vibration levels were 4g rms for level 1 and 8g rms for level II. However, the results of the initial tests conducted at 8g rms indicated that the missile failures that occurred were not typical of the failures that occur in service, in that the ratio of mechanical to electrical failures was too high. It had been believed originally that 8g rms would be an acceptable test level because it is very nearly the same as the level used by the contractor in the qualification of the missile. The occurrence of atypical failures indicated that the missiles under test could not meet the first test criterion and, although vibration testing at 8g rms might be a sound technique for demonstrating the structural adequacy of the missile and missile components, it was too severe for use in a reliability measurement test. Therefore, the test levels were changed to 2g rms for level 1 and 4g rms for level II, and the results of all previous tests at 8g rms were disregarded for MTBF considerations.

2. Excessive test delays were caused by an intermittent failure in one of the missiles (A) under test. Shortly before the conclusion of the tests, a cold-soldered joint was found on the ungrounded side of a capacitor. The cold-soldered joint may have been the intermittent failures that occurred throughout the testing of the missile. However, this point was not pursued, and testing of this missile was truncated at completion of the testing of the other missiles.

A review of the test history of missile A was conducted, and it was found that this missile had not been tested during this program at level II, or 4g rms. Therefore, an analysis was performed, comparing the test histories of all missiles at level I, to determine if the results of missile A were typical, or fit the population. A simple analysis was performed by comparing the total times for each missile at level I and the total number of verified failures.

Missile	Test Time (Minutes)	Number of Failures
A	520	8
B	1,080	1
C	1,052	5
D	1,003	2

MTBF could also have been used; however, when only one failure occurs with a large amount of test time, little confidence can be placed on the resultant calculated MTBF. From the above data, it can be seen that the total test time for missile A was approximately one-half that of any other missile, and the number of verified failures was approximately twice that for any other missile. It is obvious, without placing statistical confidence in the data, that missile A not fit the general population of the other missiles (since the MTBF would be about one-fourth of any other missile), and the missile A results for MTBF calculations were neglected.

Data Validity Investigation. Before placing much emphasis on the foregoing data it is necessary to investigate the comparison of these results to Fleet failure data. This is done by determining if the test acceleration limits were valid (if the failures in the test were typical of those reported by the Fleet) and if the MTBF data can be mathematically modeled to provide a reasonable estimate of in-service captive-flight reliability.

The determination of whether or not the acceleration of the test is too great is dependent on the comparison of the failures in test with the failure in service. A primary point of comparison is the ratio of mechanical to electrical failures. This point of comparison is based on the premise that with a missile operating in any static environment, certain electrical breakdowns are anticipated. If dynamic environments (shock and vibration) are added, some mechanical failures will occur. The difference between these failure modes is in the assumption that electrical failures are "heat effected" and mechanical failures are "non-heat effected." Although admittedly crude, this criterion was used because this is the rating criterion used for years in analyzing Fleet return failures. The failures of missiles in service can be assessed by the ratio of mechanical to electrical failures. To assess the validity of the test conditions used in the Accelerated Failure Rate Test program, the ratio of mechanical to electrical failures in the test missiles was compared to a similar ratio for missiles returned from Fleet service.

The data used for this comparison of missiles were taken from postdeployment, off-load inspections of missiles from two Navy carriers and one Marine Corps organization. The data were selected from previous investigations where the cause of missile failures had been identified and was considered typical of missile failures occurring in service. Review of these data indicates that 56 percent of the primary failures in the Fleet-returned sample were mechanical. By comparison 60 percent of the primary failures which occurred during the accelerated failure rate tests were mechanical. The comparison indicated that the stresses imposed during these tests were comparable to the stresses encountered by the Fleet-returned sample.

Acceleration Model Investigation. The total test time and the total number of failures for all missiles (neglecting missile A) were summed at each vibration level. The results obtained were:

Vibration Amplitude (g Rms)	Total Test Time (Minutes)	Number of Failures	MTBF (Minutes)
2	3,135.3	8	391.93
4	780.6	12	65.06

This analysis indicates that doubling the vibration level from 2 to 4g rms decreases the MTBF, or accelerates the test, by a factor of 6.02.

An earlier section of this paper reviewed a theory relating vibration level to time to failure. Briefly the theory is

$$\frac{T_2}{T_1} = \left(\frac{g_1}{g_2}\right)^{2.25}$$

where T is the time to failure and g is the rms vibration level. Assuming that time to failure is directly proportional to the MTBF measured in an accelerated failure rate program, the theory can be extended to the following

$$\frac{MTBF_2}{MTBF_1} = \left(\frac{g_1}{g_2}\right)^{2.25}$$

Thus, doubling the vibration level should decrease the MTBF by a factor of 4.76.

From the MTBF data of this program, a calculation can be performed to determine the power obtained during the test program. The calculations are:

$$\left(\frac{4}{2}\right)^X = \frac{391.93}{65.06}$$

and

$$X = 2.59$$

Thus, doubling the vibration level will decrease the MTBF by a factor of 6.02 or 27 percent greater than previously suggested. This difference is not too large and can be accounted for in experimental error or variations in the material coefficients b and n that define the power.

The vibration levels used in these tests were derived from measurement of severe captive-flight conditions—conditions designed to accelerate failure rate from that which would occur during normal captive flight. Very little published data exist to define the vibration levels of this particular missile during normal captive flight; however, an estimate was made based upon a small amount of unpublished data and a narrow range of possible values defined. Based upon the model determined from experimental results, a prediction of MTBF can be made for the end points of the estimated range and, assuming exponential distribution theory, a prediction of Fleet reliability can be made. When this was done, it was found that predicted reliability for normal conditions compared within a couple of percent of that reported by the Fleet, thus providing further evidence of the accuracy of the model and the validity of these tests.

Discussion

Upon reviewing the results of the foregoing tests, two factors became evident: (1) the model suggested by Curtis et al., using a power of 2.25, is reasonably accurate, and (2) there is a definite limit to valid acceleration.

The foregoing test program gave results indicating an acceleration factor of 2.59 as opposed to 2.25. Within the limits of experimental error, these figures represent close agreement. Although the standard deviation for the acceleration factor of the test program was not calculated and the individual failure time records are not available, it is believed that the calculated acceleration factor minus one standard deviation, or the $2.59 - \sigma$, would encompass 2.25, thus demonstrating statistical agreement. In addition, it must be remembered that 2.25 is not an absolute number. It was calculated from an estimate of typical material properties and may, for any system, be slightly larger or smaller.

Initially, the vibration levels had been chosen as 4 and 8g rms. Shortly after initiation of testing at 8g rms, it became obvious that failures were occurring in the missile that were not typical of those reported in Fleet return data. An investigation of the ratio of mechanical failures to electrical failures at this level revealed that this ratio had drastically changed from the ideal. This demonstrated that acceleration to 8g rms was not valid and, correspondingly, a limit of validity existed for this missile somewhere between 4 and 8g rms.

Conclusion

Based upon these analyses and test results, it is concluded that, for most air-to-air guided missiles, an acceleration factor between 2.25 and 3.00 is reasonable and should provide accurate data. It is also concluded that a limit for acceleration exists, as would be expected from S-N considerations, and in any accelerated test program, care should be taken to verify that this limit is not exceeded.

Reference

1. Shock and Vibration Information Center Monograph No. SVM-8, "Selection and Performance of Vibration Tests," by Allen J. Curtis, Nickolas G. Tinling, and Henry T. Abstein, Jr., 1971, U.S. Naval Research Laboratory, Library of Congress Catalog No. 71-176236.

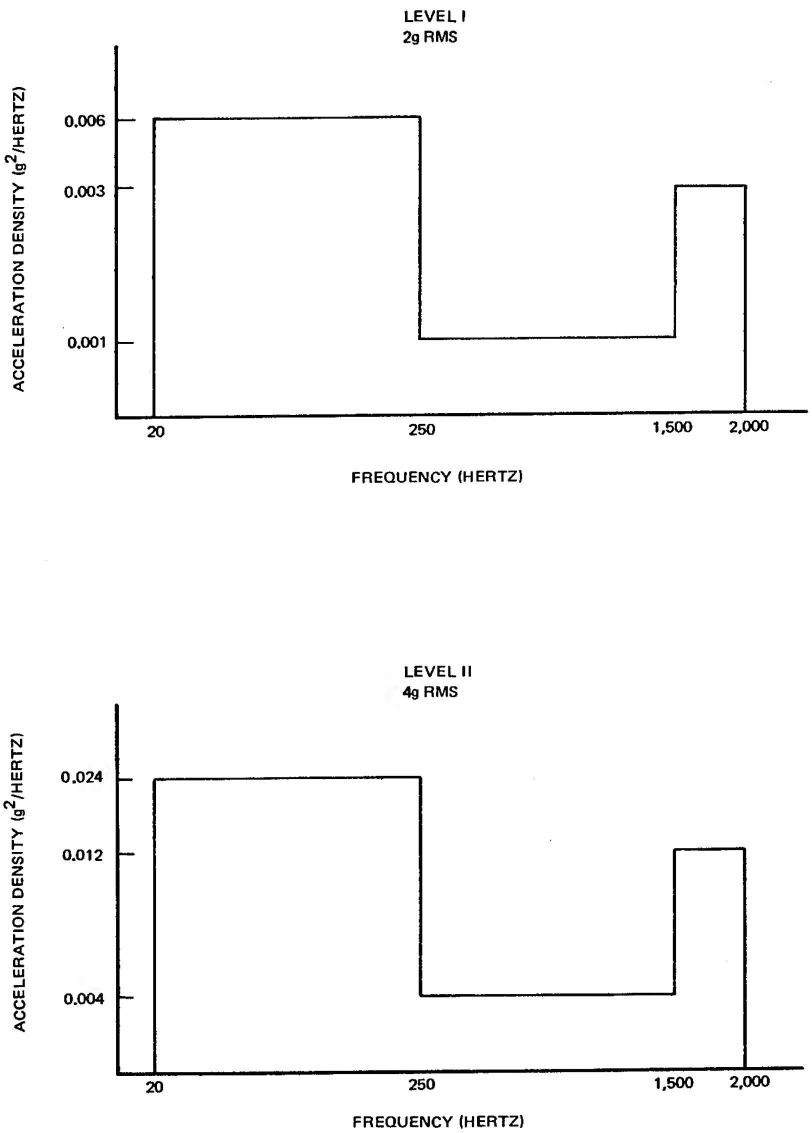


Figure 1. Vibration Acceleration Spectral Densities.

FAILURE RATE FUNCTIONS FROM TEST DATA

Henrik V. J. Broberg
Inger A. M. Gustafson
S. G. Fredrik Sandin

Military Electronics Laboratory (FTL)
Stockholm, Sweden

Summary

Presented in this paper is a computerized method for analysis of reported data from operation and testing of components and equipment. The system is based on the hazard plotting technique.

When treating the information, consideration is taken for every tested unit to the following factors: age and test time, failure mode, failure definitions, operational and environmental parameters. As an intermediate result the mean failure rate and confidence limits for times between consecutive observations are given and plotted. Owing to the pattern of the failure rate functions, the operator may choose different statistical distributions for three different intervals of age, so that continuous functions for failure rate and confidence limits can be given.

Introduction

In this paper the interest will be focused on reliability data and the manipulations necessary in order to make them useful. Figure 1 illustrates this and indicates, that for the moment the lack of relevant data may be thought of generally as a narrow sector in reliability engineering.

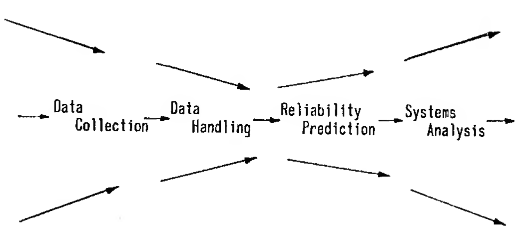


Figure 1

From the view of the systems analyst, reliability data is only one among a number of in-puts to the systems models, as is seen from figure 2.

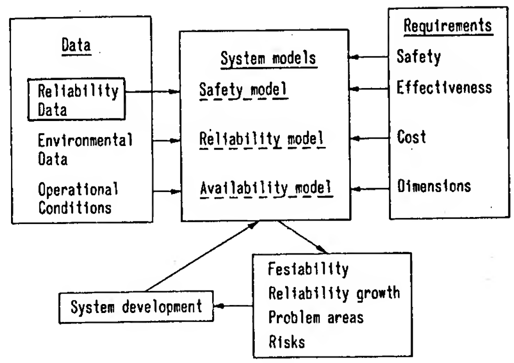


Figure 2. "Reliability data" in connection to the process of system development

From requirements on safety, effectiveness, cost dimensions and data concerning reliability, environment and operating conditions, conclusions are taken after use of the systems models. Very often rough guesses on reliability data are used in this process, which indeed violates the result.

When treating the reliability concept on the level of components or equipment units, the picture may be the one of figure 3.

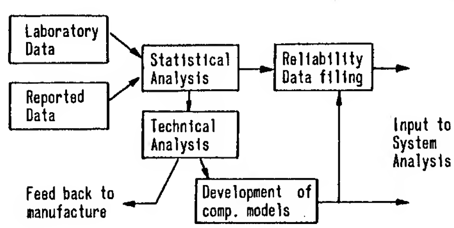


Figure 3. Handling of reliability data

A stream of laboratory and reported data is entering the office of the reliability engineer. He must analyse the information statistically and technically and feed his findings back to manufacture, if this is the aim of his job. If he is serving a systems analysis group, then he must file data for future use. As a research task he may develop models of the statistical behaviour for some types of reported or tested units.

In the interface between component reliability engineering and systems analysis we find the problem of data acquisition. For a specific analysis of a system there will generally not be much time to grasp the reliability information. A data base with a flixible output unit will therefore be necessary if reliability data shall come to practical use. See figure 4.

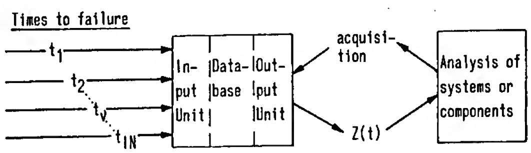


Figure 4. Database for reliability data.

In the data base we now assume, that for each registred unit a number of reported cases are stored. Each case is decribed by information on environment, operating conditions, failure modes and failure definitions. Also all times to end of test are stored for each unit. Every such time is connected to a failure mode and definition. If reported time is ended by other reason then failure, this is mentioned.

When taking data out of the bank, we first assume that at least one suitable case is found. For the valid failure definitions a number of times to failure are given by the bank as well as times to other reasons for taking the units out of test.

Starting with those unit-times we want to estimate the failure rate as a function of life-time. The basic method chosen for this is here the hazard plotting technique, described by Wayne Nelson in Proceedings from 1969 Annual Symposium on Reliability.

At FTL in Stockholm a computer program has been developed, where methods based on this technique are used for estimating failure rates with confidence limits concerning all three periods in the life of a unit: "burn in", "normal life" and "wear out". These methods will be described in the following.

The hazard function

Reliability and failure rate from the hazard function

Starting with the formula

$$R(t) = e^{-\int_0^t Z(x)dx} = e^{-H(t)}$$

where H(t) is the so called hazard function, we have the failure rate

$$Z(t) = \frac{dH(t)}{dt}$$

Suppose, that N units have been tested (or reported) and the times to failure, t_i, has been observed. The times are ordered in size:

$$0 < t_1 < t_2 < \dots t_i < \dots < t_N$$

$$\Delta t_i = t_i - t_{i-1}$$

In every interval between failures we associate the failure rate function with the mean failure rate in the interval. Now we approximate the total failure rate function with constant steps, equal to the mean failure rates in the intervals.

If the failure rate is constant in an interval, then the number of failures is proportional to the length of the interval:

$$Z_i \Delta t_i = \frac{\Delta N}{N_i}$$

Let us now define $h_i = Z_i \cdot \Delta t_i$. Only the case $\Delta N = 1$ failure is assumed.

Hence $h_i = \frac{1}{N_i}$

The connection between the hazard function, H_i and h_i is:

$$H_i = \int_0^{t_1} Z(x)dx = \sum_{\nu=1}^{i} Z_\nu \Delta t_\nu = \sum_{\nu=1}^{i} h_\nu$$

By estimating $\{h_\nu\}_{\nu=0}^{i}$ also H_i is estimated as the sum of all h_ν.

For a sample of N reported times to failure we find

$$\begin{cases} h_1 = \frac{1}{N} \\ t = t_1 \end{cases} \qquad \begin{cases} h_2 = \frac{1}{N-1} \\ t = t_2 \end{cases}$$

$$\begin{cases} h_\nu = \frac{1}{N-\nu+1} \\ t = t_\nu \end{cases} \qquad \begin{cases} h_N = 1 \\ t = t_N \end{cases}$$

For a time interval (t_{i-1}, t_i) the mean failure rate can be estimated as

$$Z(t_i, t_{i-1}) = \frac{H(t_i) - H(t_{i-1})}{t_i - t_{i-1}} = \frac{h_i}{\Delta t_i}$$

If the failure rate does not follow an exponential distribution it is, however, possible to estimate the mean failure rate for each interval between failures. As an example on this we may study the following case.

Example

Ten units have been life tested. The time to failure for each unit has been observed. In figure 5 is a table given, containing the times and calculated values for h and H. In figure 6 corresponding diagrams of failure rate and H-function are given. The data in the example are generated from a random-table and are intended to follow a life distribution with the constant failure rate $Z = 10.0 \cdot 10^{-3}$ f/h.

Failure nr (i)	t_i	Δt_i	h_i	H_i	Z_i	
0	0	0	0	0	0	
1	12	12	0.10	0.10	8.3	10^{-3}
2	29	17	0.11	0.21	6.5	"
3	39	10	0.13	0.34	12.5	"
4	70	31	0.14	0.48	4.6	"
5	87	17	0.17	0.65	9.8	"
6	106	19	0.20	0.85	10.5	"
7	134	28	0.25	1.10	8.9	"
8	173	30	0.33	1.43	8.5	"
9	218	45	0.50	1.93	11.1	"
10	301	83	1.00	2.93	12.0	"

Figure 5 Example of data analysis with the hazard-plotting method. A number of ten units are tested to failure.

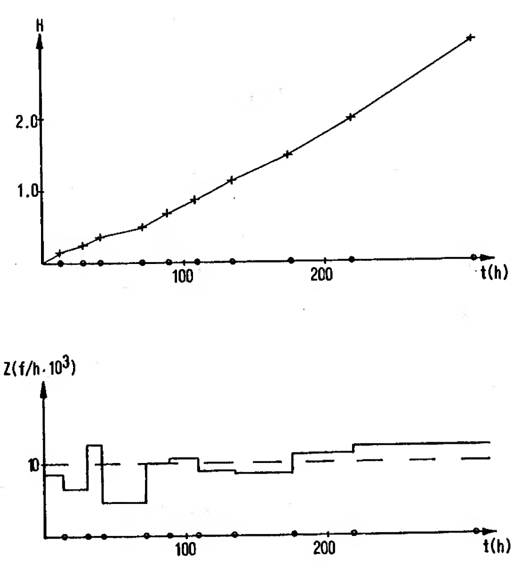


Figure 6. Example of an H-plot and the associated failure rate.

The in this manner generated partly constant failure rate curve gives a detailed picture of the reported information. The diagram, however, can show heavy oscillations, although the "real" life distribution has a continuously varying failure rate. This is because of the randomness of the failure occurrence.

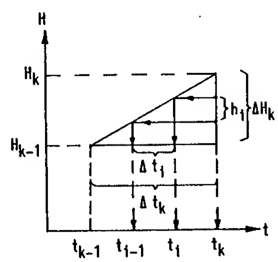
Numerical changes in the population

It is not always possible to observe the exact moment when a failure occurres. The observations may be performed at administratively suitable times. The observed number of failures will probably not have occurred all in the end of the interval. Correction for this may be done by distributing the failures in the interval.

The number of units in the population can change not only because of failures, but also by other reasons. When units are taken out of test between failures, corrections must be done for the decrease of test time.

Intervals with more than one failure

When more than one failure has occurred in the interval between two consecutive observations, the times to failure can be spread into the interval as expected times to first, second etc, failure from an exponential distribution. This is easily done by assuming the hazard function to be linear in the interval as is illustrated in figure 7.



$$t_i = t_{i-1} + \frac{\Delta t_k}{\Delta H_k} \cdot h_i$$

Figure 7. Exponential spreading out of times to failures in an interval.

This spreading out of failures is practical, as it then is possible to concentrate on intervals with only one failure. When calculating confidence limits, it has also been found that this approach will give a very nice result. It will also follow, that the later described "smoothing method" applied on the steps in the interval will modify the individual failure rate levels to be more continuously attached to the surrounding intervals.

An example of a case where the population was observed once a year is given in figure 8.

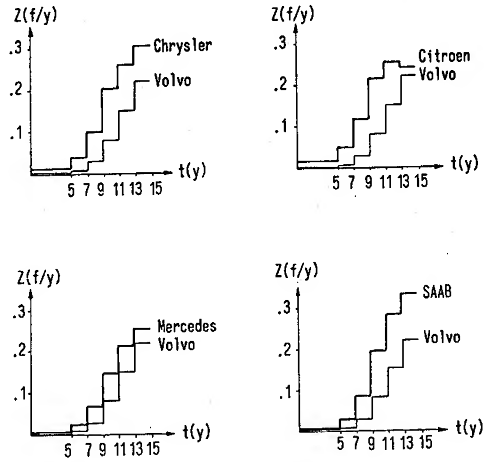


Figure 8. Example of the step-curve for failure rate. Different cars "scrap-rate" is plotted from Swedish statistics. The cars are manufactured ca 1955.

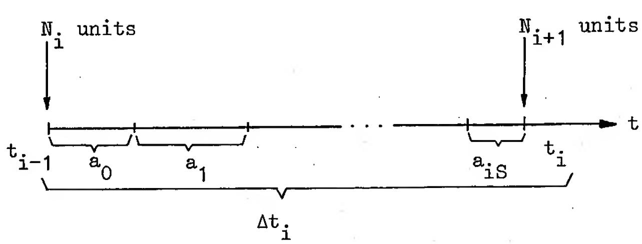
<u>Units taken out of test by other reasons than failure</u>

For example, failures in other modes than those relevant for the actual analysis will not be treated as failures. The units taken out by other reason than failure will, however, produce test time, which must be taken into account.

Suppose, that the number of units in the test is N_i in the beginning of the interval

$$\Delta t_i = (t_{i-1}, t_i).$$

The interval contains one failure in the end and S_i cases of units taken out of test by other reasons in the end of the times $a_{i0}, a_{i1} \dots a_{iS-1}$



The failure rate is assumed to be constant in the interval. Then we have

$$\begin{cases} Z_i\, N_i \cdot a_{i0} & = 0 \\ Z_i\, (N_i-1)\, a_{i1} & = 0 \\ Z_i\, (N_i-2)\, a_{i2} & = 0 \\ \text{-------------} & \text{---------} \\ Z_i\, (N_i - S_i + 1)\, a_{iS-1} & = 0 \\ Z_i\, (N_i - S_i)\, a_{iS} & = 1 \end{cases}$$

$$\begin{cases} Z_i N_i a_{i0} + Z_i (N_i-1) a_{i1} + \dots + Z_i (N_i-S) a_{iS} = 1 \\ Z_i N_{xi} \Delta t_i = 1 \end{cases}$$

where N_{xi} is an eqvivalent of the number N_i modified for the units taken out of test without failure.

Hence

$$\begin{cases} N_{xi} = N_i - \dfrac{1}{\Delta t_i} \displaystyle\sum_{\nu=1}^{S_i} \nu\, a_{i\nu} \\ h_i = \dfrac{1}{N_x} \\ N_{i+1} = N_i - 1 - S_i \end{cases}$$

When N_i has been modified to N_{xi}, the interval may be treated as an interval where only one failure has reduced the number of units. The method for spreading out failures, given in figure 8, must, however, be applied before the N_{xi}:s are calculated.

<u>Filtering the observed failure rate</u>

In order to filter statistical noise from the failure rate function a method has been developed to smooth the curves. The principle is that information from surrounding intervals will influence the failure rate in each step.

In order to influence the failure rate in the interval (t_{i-1}, t_i) from the surrounding intervals, we will use the slopes from the segments in the H-plot.

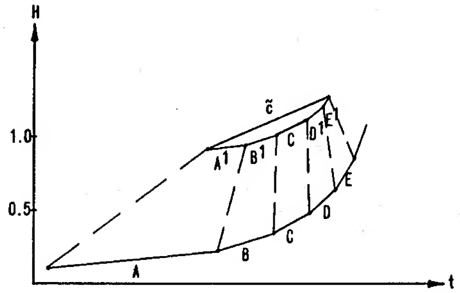


Figure 9. Recalculation of the slope of an interval (C) with the smoothing method.

The simple basis for the chosen technique may be seen from figure 9. Suppose we want to replace the slope of segment C with a figure influenced of the surrounding segments. Move C in parallel to C¹. From segments D and B we move fractions of each in parallel and connect them on each side to C', thereby getting B' and D'. Then we move smaller fractions of A and F and connect them to B' and E'. Finally connect the ends of A' and E' with a straight line. The slope of this line will now be our replacement for the slope of segment C.

For the procedure of influncing one interval from the others, two axioms will be used:

A) The weight of the influence shall not be laterally biassed.
B) The nearest intervals shall have the strongest influence.

When calculating the failure rate Z_i for the interval (t_{i-1}, t_i) we will take care of the information from surrounding intervals by introducing two series of coefficients $\{\alpha_\nu\}$ and $\{\beta_\nu\}$:

$$\tilde{Z}(\Delta t_1) = \frac{\sum\limits_{\nu=Max(0,i-n)}^{i-1} \alpha_\nu \beta_\nu h_\nu + h_i + \sum\limits_{\nu=i+1}^{Min(N,i+n)} \alpha_\nu \beta_\nu h_\nu}{\sum\limits_{\nu=Max(0,i-n)}^{i-1} \alpha_\nu \beta_\nu \Delta t_\nu + \Delta t_i + \sum\limits_{\nu=i+1}^{Min(N,i+n)} \alpha_\nu \beta_\nu \Delta t_\nu}$$

where n gives the maximum number of terms on each side of the interval in question and $\tilde{Z}_i$ is the smoothed failure rate. Now let α_ν take care of axiom A and β_ν of axiom B. For $\{\alpha_\nu\}$ we have the series:

$$\alpha_\nu = \frac{N-\nu+1}{N-i+1}$$

which is slowing down with the same rate as h_ν is growing and fills the condition $\alpha_i = 1$.

For $\{\beta_\nu\}$ a geometric series is convenient to choose:

$$\beta_\nu = k^{|t-\nu|}, \quad 0 < k < 1$$

For $\{h_\nu\}$ we have

$$h_\nu = \frac{1}{N-\nu+1}$$

The term $\alpha_\nu \beta_\nu h_\nu$ from the numerator of the failure rate formula will get the value:

$$\alpha_\nu \beta_\nu h_\nu = k^{|i-\nu|} \cdot h_i$$

The formula for Z_i can then be rewritten as

$$\tilde{Z}(\Delta t_i) = \frac{h_i \sum\limits_{\nu=Max(0,i-n)}^{Min(N,i+n)} k^{|i-\nu|}}{\sum\limits_{\nu=Max(0,i-n)}^{Min(N,i+n)} k^{|i-\nu|} \frac{N-\nu+1}{N-i+1} \cdot \Delta t_\nu}$$

Where the sum in the numerator is simply evaluated as

$$\sum_{\nu=Max(0,i-n)}^{Min(N,i+n)} k^{|i-\nu|} = \frac{1+K}{1-K} - \frac{K}{1-K} \left(K^{Min(n,i-1)} + K^{Min(n,N-i)}\right)$$

Different smoothing formulas for the calculation of failure rate can be assumed by choosing different geometric series $\{\beta_\nu\}$ and varying the number n. By using random numbers generators and a computer it is possible to study how the formulas will work on simulated times to failure from known distributions. Such work has been done and will be a basis for further development.

In figure 10 an example is given on the smoothing technique. Three curves are given for comparison. All of them are based on the data in figure 5. The upper curve is not smoothed. The middle one is for each interval taking influence from one neighbour on each side and the lower takes two neighbours from each side into account. The factor in the geometric series is 0.5.

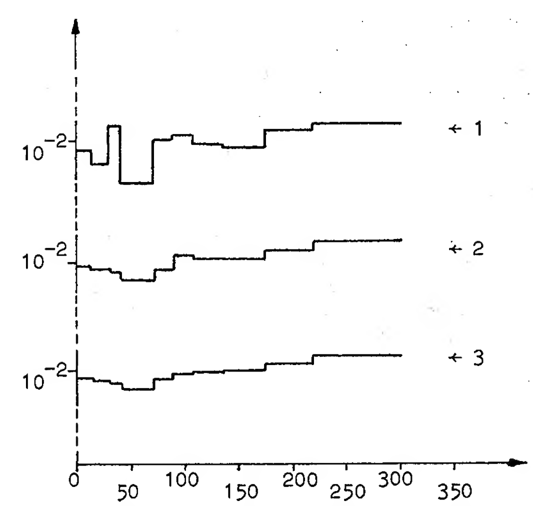


Figure 10. Example of the smoothing method.

As is seen from the figure, the filtering effect of the method is obvious.

Confidence limits

In order to get some knowledge of the confidence in the estimates of failure rate we have made an approach to the calculation of confidence levels for the H-function. Those levels are then used for computing confidence levels for the failure rate function. The basis for the calculations is mainly the assumption that the failure rate in each time interval can be associated with a constant failure rate equal to the mean failure rate in the interval.

Confidence limits for time between failures

From the literature it is known that the mean failure rate, for a time interval ended when r failures have occurred, can be assumed to follow a χ^2-distribution with 2r degrees of freedom.

The upper and lower p % confidence limits for the mean failure rate are expressed by

$$Z_i^* \frac{\chi^2_{1-p/2}(2)}{2r} < Z_i < Z_i^* \frac{\chi^2_{p/2}(2r)}{2r}$$

where $Z_i^* = r/\Delta t_i$ is the observed failure rate. In our case we have only one failure and the length of the time-interval is therefore the time to first failure. Hence, the confidece limits for the length of the time-interval are

$$\begin{cases} \text{Upper limit } \overline{\Delta t}_i = \Delta t_i^* \cdot \dfrac{2}{\chi^2_{1-p/2}(2)} \\ \text{Lower limit } \underline{\Delta t}_i = \Delta t_i^* \cdot \dfrac{2}{\chi^2_{p/2}(2)} \end{cases}$$

where Δt_i^* is the observed length of time.

Confidence limits to the hazard function

Suppose that we are studying the growth of the H-function in the i:th interval. We assume that during all of the interval H is growing linearily from H_{i-1} to H_i. Now we introduce the function $h_i(t)$, which is linear and grows from $h_i(0) = 0$ to $h_i(\Delta t_i) = \frac{1}{N-i+1} = h_i$, which is a constant established for each interval.

It can, however, be argued on the length Δt_i of the interval. From the formulas above, upper and lower limits for Δt_i are given. If the upper limit $\overline{\Delta t}_i$ is taken as an estimate of the interval, then for the observed time:

$$h_i(\Delta t_i^*) < h_i(\overline{\Delta t}_i) = h_i$$

On the contrary, if the lower limit is used as an estimate, then

$$h_i(\Delta t_i^*) < h_i(\underline{\Delta t}_i) = h_i$$

The function h(t) can therefore be assumed to have a slope distributed with an upper and a lower limit. Hence also $h_i(\Delta t_i^*)$ for the observed length of time Δt_i^* will be a statistical variable, for which confidence limits can be established. For each interval we now introduce the statistical variable

$$h_i^* = h_i(\Delta t_i^*)$$

In figure 11 two functions for $h_i(t)$ are given.

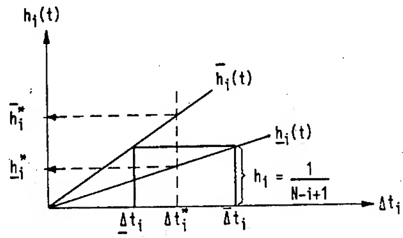


Figure 11. Confidence limits for the statistical variable h_i^*

One upper function $\overline{h}_i(t)$ is passing the point $(h_i, \underline{\Delta t}_i)$ and a lower function the point $(h_i, \overline{\Delta t}_i)$. From homogeneous triangels in the figure it is easy to find the formulas.

$$\begin{cases} \overline{h}_i = \dfrac{t_i^*}{\underline{\Delta t}_i} \\ \underline{h}_i = h_i \dfrac{t_i^*}{\overline{\Delta t}_i} \end{cases}$$

If we substitute $\underline{\Delta t}_i$ and $\overline{\Delta t}_i$ with their χ^2-equivalents, the upper and lower limits for h_i^* are given by:

$$\text{Upper limit } \overline{h}_i^* = \frac{1}{N-i+1} \cdot \frac{\chi^2_{p/2}(2)}{2}$$

$$\text{Lower limit } \underline{h}_i^* = \frac{1}{N-i+1} \cdot \frac{\chi^2_{1-p/2}(2)}{2}$$

Hence, for an arbitrary time interval Δt_i, the expression $x = 2(N-i+1)h_i^*$ is χ^2-distributed with 2 degrees of freedom. The differential of the distribution for x is:

$$f(x)dx = \frac{1}{2} e^{-\frac{x}{2}} dx$$

By the variable transformation $x = 2(N-i+1)h_i^*$ we achieve

$$f(h_i^*)dh_i^* = (N-i+1)\ e^{-(N-i+1)h_i^*}\ dh_i^*$$

The H-function is built up by summing h-functions. The confidence limits for H_i is therefore depending on information from earlier intervals. In order to establish those intervals we are studying the distribution for the sum:

$$H_i^* = \sum_{\nu=1}^{i} h_\nu^*$$

It is possible to evaluate this distribution, but we have chosen to use an approximation, which will give a good fit to the exact distribution. This approximation is based on the well known use of a normal distribution and its derivates. The approximate frequency function for H_i is

$$\begin{cases} f(x) = \varphi(x) - \dfrac{\gamma_1}{6}\varphi^{III}(x) + \dfrac{\gamma_2}{24}\varphi^{IV}(x) + \\ \qquad + \dfrac{\gamma_1}{72}\varphi^{VI}(x) \\ x = \dfrac{H_i^* - m}{\sigma} \end{cases}$$

where $\varphi(x)$ = the frequency function for the normal distribution
x = normalized variable
γ_1 = obliquity for H_i^*
γ_2 = excess for H_i^*

The parameters γ_1 and γ_2 are computed from the generic function for H_i^*, which is the product of the generic functions for h_i^*.

For h_i^* the generic function is:

$$\psi(t)_{h_i^*} = \frac{N-i+1}{N-i+1-t}$$

Hence, the generic function for H_i^* is

$$\psi(t)_{H_i^*} = \prod_{\nu=1}^{i} \frac{N-\nu+1}{N-\nu+1-t}$$

From this function the following is derived:

1) Mean value

$$m = \sum_{\nu=1}^{i} \frac{1}{N-\nu+1} \longrightarrow H_i$$

2) variance

$$\sigma^2 = \sum_{\nu=1}^{i} \frac{1}{(N-\nu+1)^2}$$

3) obliquity

$$\gamma_1 = \frac{1}{\sigma^3} \sum_{\nu=1}^{i} \frac{2}{(N-\nu+1)^3}$$

4) excess

$$\gamma_2 = \frac{1}{\sigma^4} \sum_{\nu=1}^{i} \frac{9}{(N-\nu+1)^4} - 3$$

The lower x % confidence limit for H_i^* is then found from

$$\begin{cases} \displaystyle\int_{-\infty}^{x} f(x)dx = \frac{x}{100} \\ \underline{H}_i^* = x \cdot \sigma + M \end{cases}$$

The upper $(1 - \alpha)$ % confidence limit for H_i^* is in the same manner found from

$$\begin{cases} \displaystyle\int_{-\infty}^{x} f(x)dx = \frac{1-x}{100} \\ \overline{H}_i^* = x \cdot \sigma + m \end{cases}$$

<u>Confidence limits for failure rate</u>

Confidence limits for the failure rate is easily found from the confidence limits for the H-function.

Suppose, that we for a component have three hazard functions; one upper p % function $\overline{H}(t)$, one mean function $H(t)$ and one lower p % function $\underline{H}(t)$. See figure 12.

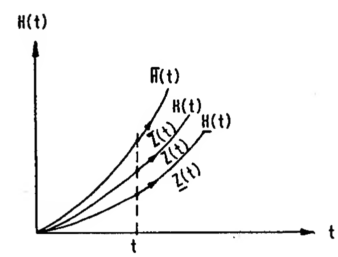


Figure 12. Upper and lower limits for the H-function and the corresponding failure-rates at the time t.

The mean failure rate for the time interval (0, t) is

$$Z_m(t) = H(t)/t$$

From the upper and lower H-functions the corresponding functions for the mean failure rate are found

$$\begin{cases} \text{Upper limit} \quad \overline{Z}_m(t) = \dfrac{\overline{H}(t)}{t} \\ \\ \text{Lower limit} \quad \underline{Z}_m(t) = \dfrac{\underline{H}(t)}{t} \end{cases}$$

The relationship between $Z_m(t)$ and $Z(t)$ is:

$$Z_m(t) = \frac{1}{t} \int_0^t Z(x)dx$$

The upper limit for the mean failure rate can then be connected to an upper failure rate function

$$\overline{Z}_m(t) = \frac{1}{t} \int_0^t \overline{Z}(x)dx$$

The upper H-function is expressed by

$$\overline{H}(t) = t\, \overline{Z}_m(t) = \int_0^t \overline{Z}(x)dx$$

Hence,

$$\frac{d\overline{H}(t)}{dt} = \overline{Z}(t)$$

For each interval (t_{i-1}, t_i), in which the H-function is assumed to be linear, the upper p % limit for the failure rate is

$$\overline{Z}_i = \frac{\overline{H}_i - \overline{H}_{i-1}}{t_i - t_{i-1}} = \frac{\Delta\overline{H}_i}{\Delta t_i}$$

The lower limit for the interval is

$$\underline{Z}_i = \frac{\underline{H}_i - \underline{H}_{i-1}}{t_i - t_{i-1}} = \frac{\Delta\underline{H}_i}{\Delta t_i}$$

Z_i as well as $\overline{Z}_i$ and $\underline{Z}_i$ are plotted by the computer. The operator may choose the smoothing parameters he wants in order to study the reliability trend for the tested (or reported) units in various parts of the life curve. The same smoothing is used for the confidence limits as for the mean estimate. How this influence the limits is not generally possible to evaluate and will not be discussed here.

Example

An example of the resulting curves is given in figure 13. A number of 230 polyester capacitors were tested for 10,000 hours. The failure definition was "short-circuit caused by dielectric break down". An overstress of double rated voltage was applied and the temperature was 85°C.

The first observation was at 100 h, when 31 failures already had occurred. In the calculations constant failure rate is assumed for this interval. The exponential distribution of the failures gives wide confidence limits in the start and more narrow, when a growing number of failures is taken into account.

Although the scale on the Z-axis is logarithmic, it is obvious that some kind of a bath-tube curve is found. Only a slight smoothing is performed and the curve looks a little noisy. But there is probably a physical background to the heavier variations. The plot could therefore give indications on different aging processes going on.

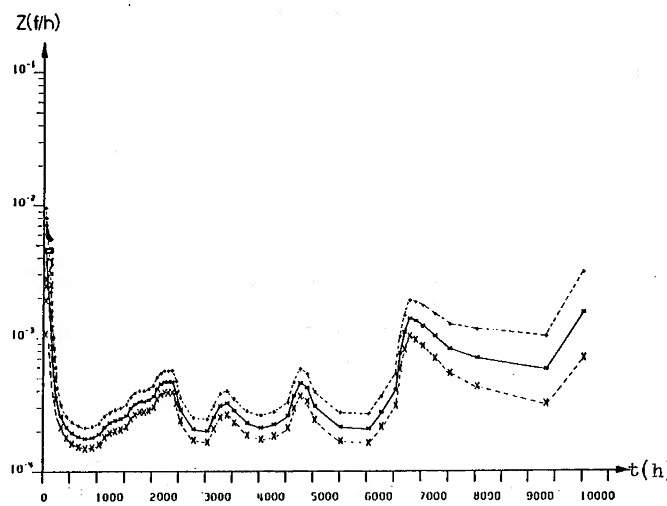
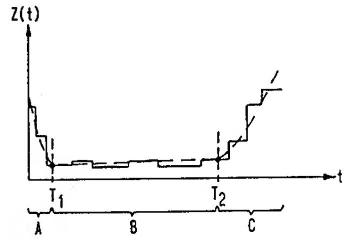


Figure 13. Step-approximations of failure rate from a reliability test of a polyester capacitor (FRD-card No 136). Upper and lower 90% confidence limits are marked with dotted lines.

Curve fitting

To the failure rate step functions are statistical functions fitted and matched together. In figure 14 are the functions and their regions given. The fitting is made with a least square method applied on the step functions for the failure rate with upper and lower confidence limits.



A) Early failure period

$$Z_A(t) = a \cdot e^{bt}$$

B) Normal life period

$$Z_B(t) = c + dt$$

C) Wear-out failure period

$$Z_C(t) = f \cdot t^g$$

Figure 14. Curve-matching to the Z_i-function.

The two points T_1 and T_2 on the time-axis must be chosen by the analyzer after study the step-curve. Then the functions $Z_A(t)$, $Z_B(t)$ and $Z_C(t)$ will be called for to the three regions of the axis. It is possible to allow almost any choice of functions, but it is more practical to standardize the function for each region. In the FTL-program, the mean square-method is first applied on the B-region. Then the A- and C-regions are treated from the condition that their functions must match the function in the B-region.

After the equations for the curves have been calculated, their parameters can be stored and used as input data to reliability prediction programs on higher systems levels.

Applications

The methods presented here are included in a computer program at FTL. The operator is analyzing his statistical material in two or more steps. First he chooses appropriate failure modes and feeds the computer with figures on failure-times and times until units are taken out of test by other reasons. Then he will get a plot of the failure rate step-function with confidence limits. If he wants to, he can now use the smoothing technique and filter the information from random noise. When the appropriate smoothing constants are found, the next step is to choose the two times T_1 and T_2 to split the time-axis in three intervals for curve-fitting. Next print-out will give the curve equations numerically and as a plot.

The communication with the program in a future situation, where it serves as an output facility for a data-base is illustrated in figure 15.

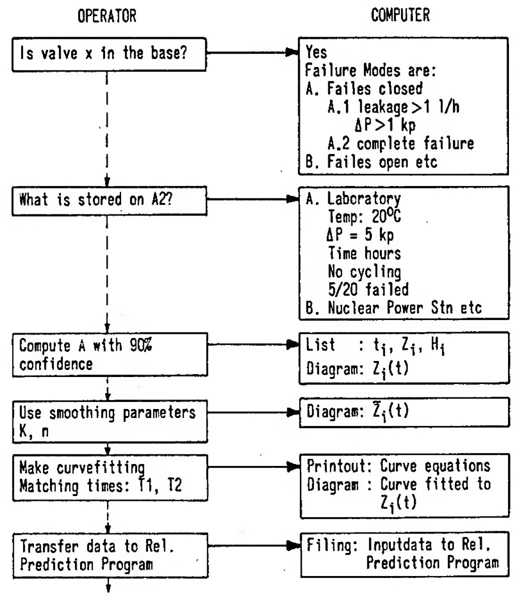


Figure 15. Example of dialog with Data Base Output Unit

The data analysis activities are of course only a part of the general reliability work. This can be examplified by a rapid view on the growing "arsenal" of computerized reliability tools at FTL. As is seen in figure 16, the intention at FTL is to cover the following areas with automated methods:

Data normalization

The program, which is not prepared yet, will automatically produce Failure Rate Data cards on components and serve as an input unit for the data-base.

Data base

For the moment a small data-base is built into the prediction program for electronic devices, RPP-1. When performing a prediction, data is automatically searched for in the bank, which is arranged in five hierarchic levels covering different identification distinctions for the components. In the future, a more general data base will be used, in which the hierarchic structure probably will be similar to the one in the existing bank.

Data Analysis

The here discussed program, DAP-1 will serve as an output unit to the Data Base and also as an input to Reliability Prediction and Systems Analysis. It can also be used separately as a statistical tool during Project Development in connection to Reliability Testing. In this case the program is practical, when analyzing the

consequences of elimination of failure modes. Also extrapolations outside the time interval for which failure information is established can be done to some degree. Another field of application is routine check-up of reported data.

Reliability Prediction

Here is assumed that Reliability Prediction mainly includes the calculation of one reliability block where all components statistically are connected in series. The program RPP-1 mainly adds failure rates for components and computes confidence limits for the sum. The content of the program is to a great extent models for the statistical behaviour of different components types in different applications and environments.

System Analysis

Figures on the reliability block level is the input to the systems analysis program RAP-1. This program is able to take almost any system structure and allows any unit to be represented in more than one function on the same time. The calculations are based on monte-carlo technique.

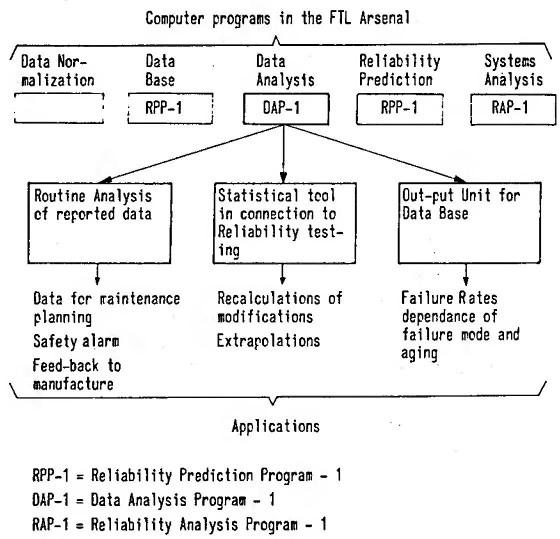


Figure 16. The frame-work of computer programs for reliability applications at FTL.

These examples of computer programs may be thought of as an example of the frame in which the in this paper described methods are supposed to work. At large, the applications for reliability data analysis techniques are as many as the applications for failure rates. As the need for accurate reliability information grows, also the data handling methods will be more important.

Conclusion

Methods based on estimation of the hazard function seems to be very promising for manipulating reliability data. The greatest power lies in the possibility to follow life distributions in detail. The "hazard-concept" is not free from contradictions to more traditional ways of treating reliability data. Extended comparative studies of methods for estimating failure rates are therefore motivated and could probably give new light to this field of reliability.

References

1) "Hazard Plot Analysis of Incomplete Failure Data"
Nelson, W.
1969 Annual Symposium on Reliability

2) "A computerized system for Reliability Prediction - RPP-1"
H. Broberg
I. Hansson
L. Jönsson
I. Persson
FTL A-report A16:23

MAINTENANCE PROCEDURES OF DEVICES FOR ROLLING STOCK AND ELEVATORS BASED ON FIELD DATA

INDEX SERIAL NUMBER – 1056

Takuji FUKUOKA
Mito Works, Hitachi, Ltd.
Ibaragi, Japan

SUMMARY

Recently the electronic equipment have been commonly adopted in rolling stock and elevators in Japan. Such equipment should be adopted some defferent maintenance methods comparing with the conventional mechanical and electrical equipment. By analyzing the field data of devices for rolling stock and elevators, this paper shows the variation of Shape parameter "m" among Weibull parameters that takes defferent distribution in according to the kinds of devices and failure modes. Moreover, the relation between Shape parameters and the maintenance methods are discussed, and fundamental procedures how should be such maintenance for each devices are derived.

INTRODUCTION

The control of rolling stock and elevators in Japan has rapidly incorporated electronic equipment. Having an inherent "maintenance-free" possibility, electronic apparatus should adopt a maintenance policy which differs from that of mechanical and electrical devices. And it should be quantively derived from the field data on actual equipment in operation.

It was decided to obtain Weibull parameters by analyzing the results of investigation on actual field data of electric, mechanical, and electronic apparatus for studying the fundamental maintenance policies of rolling stock and elevators.

Using Hitachi's products, the subject of this investigation was limited to devices for rolling stock and elevators. Conducted by the manufacturer for whom it was difficult to obtain field data, this investigation may have dealt with uncertain factors; however, it was an investigation made to form a part of the manufacturer's effort to improve reliability and maintainability.

CHARACTERISTICS OF DEVICES FOR ROLLING STOCK AND ELEVATORS

Requirements concering reliability. Since rolling stock and elevators are transportation facilities which handle passengers.

1. Safety is the prime requisite.

2. They must be free from breakdown in operation.

3. It is necessary to ensure prompt recovery service minimizing the down time caused by failure.

General maintenance method. Based on preventive maintenance principles to prevent breakdowns in operation, periodic inspection and repair have been performed.

Regarding electronic apparatus whose application has been rapidly expanding recently, conventional maintenance methods are generally employed, while there is apprehension concerning the necessity for other adequate methods.

Effects of breakdown. Since rolling stock and elevators are assemblies comprising various devices and parts, troubles involving individual devices and parts may have an effect on operation in various ways depending upon the extent of trouble'.

FIELD DATA ON DEVICES

Field data analyzing method

While Weibull distribution analyses were conducted on field data of products, information on trouble was given in various forms to the Hitachi as a manufacturer. The following describes a method of correlating these data.

Scope of data. Included in this scope are all breakdowns during operation or maintenance.

Population parameter and period. Devices of the same type are successively delivered in general cases, causing the number of the devices in operation to increase. The population parameter (number of devices) and a certain period of operation time are used as subjects.

Regarding the history of a device. Repairable devices were operated while undergoing partial repair. Thus, regarding troubles in the same region on the same device, only first failure was adopted as data for Weibull distribution.

Example of Weibull distribution.

The number of types of devices analyzed for investigation were 30 to 40 for mechanical, electric, and electronic apparatus respectively. Trouble data on these apparatus were plotted on a Weibull chart for each apparatus.

Examples of Weibull distribution are as shown in the following:

Mechanical equipment:	Units A through D, Fig. 1
Electric devices :	Units A through D, Fig. 2
Electronic equipment:	Units A through D, Fig. 3

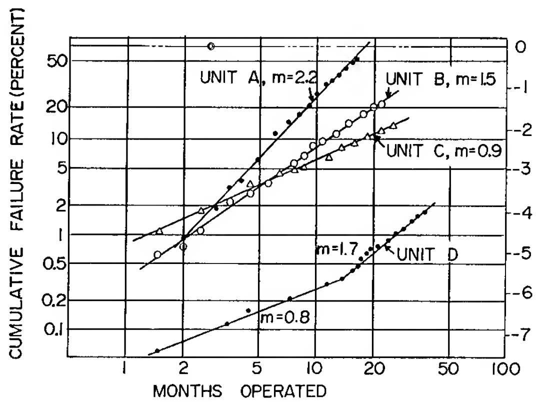


FIGURE 1. WEIBULL DISTRIBUTION IN MECANICAL EQUIPMENT

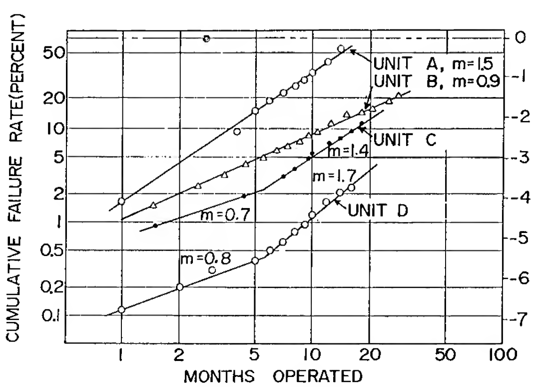


FIGURE 2. WEIBULL DISTRIBUTION IN ELECTRIC DEVICES

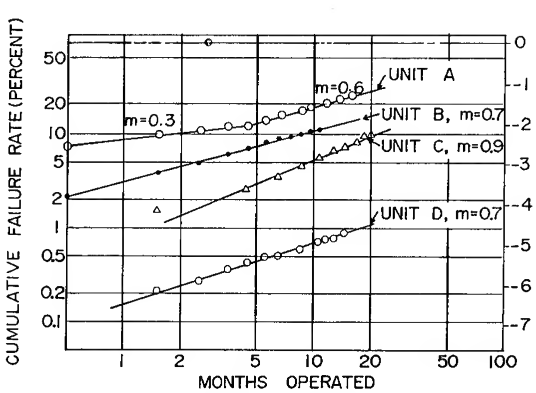


FIGURE 3. WEIBULL DISTRIBUTION IN ELECTRONIC EQUIPMENT

Shape parameter m

m for each type of device. Shape parameter m among the weibull parameter on each device was obtained by plotting on a Weibull chart for each device as shown in the examples of Weibull distribution. With apparatus roughly classfied into mechanical, electric, and electronic apparatus, distribution of shape parameter m for each kinds is shown on a log-normal chart as illustrated by Fig. 4.

1. Mechanical equipment:
   The mean value is about 1.5. Although initial failures also form a portion, wear-out failures are the largest percentage.

2. Electric equipment:
   Remarkable dispersion is noticed at m of above 2.5. Since electric apparatus contains several mechanical elements while presenting electrical troubles, initial and wear-out failures are mixed in.

3. Electronic equipment:
   m indicates a comparatively small dispersion centering around 0.7. Since a potential defect arising up during the manufacturing processes of parts or apparatus causes an unexpected failure to occur during the initial period of operation or after that, most failures occur as initial failures.

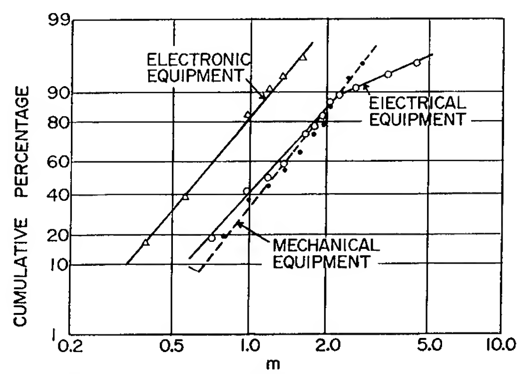


FIGURE 4. m DISTRIBUTION IN EACH TYPE OF EQUIPMENT

m obtained for each trouble region and trouble mode. With the apparatus trouble mode classified, shape parameter m was obtained by plotting on a Weibull chart for each trouble region and trouble mode. The above procedure was followed for various apparatus to obtain a number of m, which were arranged by trouble modes to be indicated on the log-normal chart. Typical distributions of m are shown by Figs. 5a, 5b and 5c.

1. Machine elements and mechanism failure
   Regarding leakage and loosening; initial failure and wear-out failure account for about half of the entire failures. There seems to be two types of causes-- (1) an improper amount of tightening and contact pressure, and dispersion in dimensions during the process of manufacture,

causing initial failure and (2) fatigue and abrasion. Especially the m distribution on loosening displays a remarkable characteristic by drastic two peak for $m < 1$ and $m > 1$ respectively.

While breakdown, cracks, and dislocation are mostly as wear-out failure pattern.

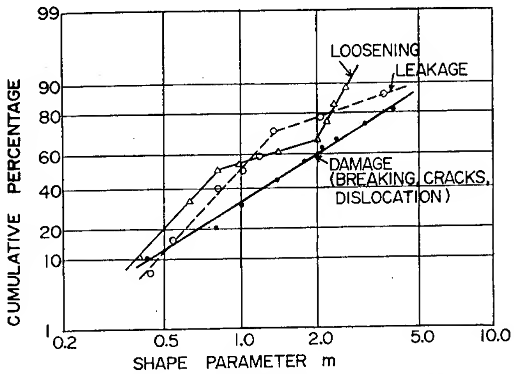


FIGURE 5a. m DISTRIBUTION IN MACHINE ELEMENTS AND MECHANISM TROUBLE

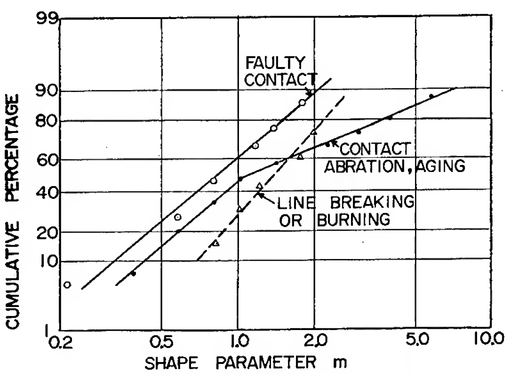


FIGURE 5b. m DISTRIBUTION IN ELECTRICAL TROUBLE

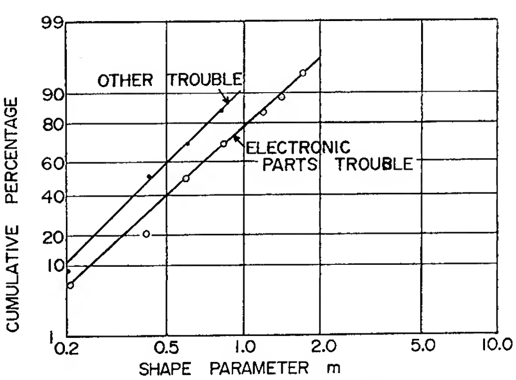


FIGURE 5c. m DISTRIBUTION IN ELECTRONIC TROUBLE

2. Electrical failure

The shape parameter on faulty contacts is below 2. Incomplete contact is found mostly in initial falure and random failure. The failures may be caused mostly by temporary incomplete contact which will not occur again.

Abrasion and aging of contacts also have an influence on incomplete contact, and input data are rarely indistinguishable each other. According to the result of study on this distribution, it is known that failures are often caused by abrasion, and there are many occasions where distribution is extended over a considerably wide range.

Regarding m on broken wires and burning, there are many cases of $m > 1$; however, dispersion is comparatively small, remaining within the range of approximately 0.8 to 2; random failure also occurs. The failure may be caused by degradation, but there is a dispersion in the progress of degradation and it takes the characteristics random time to become degraded to an unallowable failure level.

3. Electronic failure

In many cases, m on failure of electronic parts is below 1. While there are cases of $m > 1$ to the extent of approximately 20%, it may be stated that as long as ordinary parts are properly used, failures thereof may fall in the category of initial failure of $m \leq 1$ or random failure of the failure rate reduction type.

Failures other than parts failures are mostly initial failure of $m < 1$.

EXAMINATION OF MAINTENANCE PROCEDURES

Based on the results of Weibull distribution on field data of the above apparatus, examination is made on the maintenance procedures.

Shape parameter m and maintenance procedures

Reliability R(t) is generally defined as the probability that no failure will occur for a period of time t. However, in case trouble concerning a device is corrected within a period of time allowable as a system, it may not to regarded practically as a trouble. Thus, if device troubles, occured repeatedly, are recovered within allowable period of time, we obtain the practical failure rate $\lambda_E(t)$ and the total reliability $R_E(t)$ including maintenance as given by Eq.(1), (2).

$$\lambda_E(t) = e^{-\mu T} \cdot \lambda(t) \tag{1}$$

$$R_E(t) = \exp\left[-e^{-\mu T} \cdot \int_0^t \lambda(\tau)\, d\tau\right] \tag{2}$$

$\lambda(t)$: Ordinary failure rate

μ : Repair rate

T : Allowable recovery time

In case failures display a Weibull distribution, Eq. (2) becomes as follows:

$$R_E(t) = \exp\left\{-e^{-\mu T} \cdot \frac{t^m}{t_0}\right\} \tag{3}$$

Now, repeat preventive maintenance to bring the device back to a "like-new condition" at intervals of To. And in case t=nTo (n:integer) Eq.(3) becomes as follows:

$$R_E(t)=\left\{\exp\left[-e^{-\mu T}\cdot\frac{T_o^{\,m}}{t_o}\right]\right\}^{\frac{t}{T_o}}$$

$$=\exp\left[-(1-M(T))\frac{t}{t_o}\,T_o^{\,m-1}\right] \quad (4)$$

$$M(T):\ (1-e^{-\mu T})$$

Since Eq. (4) indicates a scale of reliability including corrective maintenance and preventive maintenance, such maintenance should be performed to increase the above value.

When the Weibull parameter and maintainability M(T) are known, $R_E(t)$ in Eq. (4) can be determined. Fig. 6 shows an example of how to determine the relation between To and $R_E(t)$ at a certain time (= 500hr) with m used as a parameter, while assuming scale parameter t_o and maintainability M(T).

In case m=1, $R_E(t)$ remains constant regardless of the preventive maintenance period To. In case $m>1$, $R_E(t)$ decreases accordingly as To increases, and the rate of decrease rapidly increases accordingly as m increases. In case $m<1$, $R_E(t)$ increases according as To increases and $R_E(t)$ becomes maximum when $To\rightarrow\infty$.

Therefore, in case $m\leqq 1$, $To\rightarrow\infty$; that is, preventive maintenance should not be performed; in case $m>1$, it should be performed in order to increase $R_E(t)$.

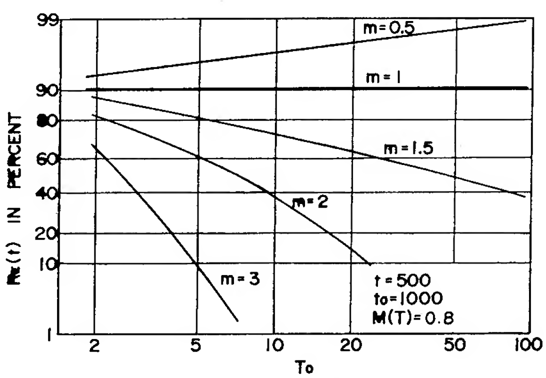


FIGURE 6. RELATION BETWEEN To AND R_E(t)

Preventive maintenance period

While it is necessary to perform preventive maintenance when parameter $m>1$, preventive maintenance requires time and expense.

The adequate period is determined according to the principle if achieving the maximum availability or minimizing maintenance costs[2,3,4] in some cases. However, it seems rational to adopt a principle of minimizing total cost consisting of actual maintenance cost and cost of loss due to down time.

When failures indicate the Weibull distribution, the average number of failures occurring during the period of To on the assumption of preventive maintenance at intervals of To is given as

$$\int_0^{T_o}\frac{m\,t^{m-1}}{t_o}\,dt=\frac{T_o^{\,m}}{t_o} \quad (5)$$

Let Tp = Mean time of preventive maintenance

T_R = Mean time of corrective maintenance

Av = Time availability

Then, unavailability A_v' for the period To is

$$A_v'=1-A_v=\frac{T_P+\frac{T_o^{\,m}}{t_o}T_R}{T_o} \quad (6)$$

$$(T_o \gg T_P)$$

On the other hand,

let Cp = Average cost of preventive maintenance

C_R = Average cost of corrective maintenance,

Then, average maintenance cost Cm for the period To is

$$C_m=\frac{C_P+\frac{T_o^{\,m}}{t_o}C_R}{T_o} \quad (7)$$

In this case, losses due to unavailability of the device vary depending on the extent of effect upon the system produced by device failure. Thus, let the loss be converted into cost by multiplying unavailability by a coefficient α which varies depending on the type of device. Then, total cost C_T is

$$C_T=C_m+\alpha A_v'$$

$$=(C_P+\alpha T_P)T_o^{-1}+(C_R+\alpha T_R)\frac{T_o^{\,m-1}}{t_o} \quad (8)$$

$$\frac{\partial C_T}{\partial T_o}=T_o^{-2}\left\{-(C_P+\alpha T_P)+(m-1)(C_R+\alpha T_R)\frac{T_o^{\,m}}{t_o}\right\} \quad (9)$$

The value of To, for which the value of Eq.(9) becomes zero, gives a maintenance period for which total cost is minimized. While the total cost cannot be minimized for $m\leqq 1$, the maintenance period of minimum total cost for $m>1$ is

$$T_{o\,min}=t_o^{\frac{1}{m}}\cdot(m-1)^{-\frac{1}{m}}\cdot\left(\frac{C_P+\alpha T_P}{C_R+\alpha T_R}\right)^{\frac{1}{m}} \quad (10)$$

As previously mentioned, certain apparatus failures may result as breakdown of operation of rolling stock or elevators, while other failures have no effect on operation.

By determining the value of α (coefficient) according to the extent of effect by failure of the subject device, an adequate $T_{o\,min}$ can be obtained by used of the Weibull parameter of field data.

Fig. 7 is a diagram used to determine $T_{o\,min}$. Obtain $(C_P+\alpha T_P/C_R+\alpha T_R)^{1/m}$ and $(m-1)^{-1/m}$ by referring to the figure and multiply to $^{1/m}$ of the Weibull distribution by the

two above values to determine the value of To min.

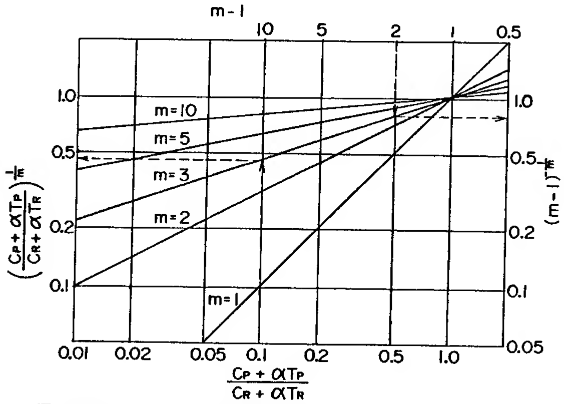


FIGURE 7. DIAGRAM OF HOW TO SEEK PREVENTIVE MAINTENANCE PERIOD

Maintenance of mechanical devices and electric devices

Preventive maintenance is effective. As previously mentioned, there are many cases of $m > 1$; thus, reliability may be improved by preventive maintenance. Even in the case of initial failures of $m < 1$, the causes are tightening, friction, and dimensions which can be detected by checking; as a result, priority is given to periodic inspections and repairs.

Preventive maintenance period and content. Table 1 is a general check list. Checking is to be made on necessary items as listed according to the device. Repairs and replacements are made on defective ones as required. While the period is to be determined by the previously mentioned adequate period, it must be determined in an actual case according to field data of the device classified by failure region and failure mode. Thus, even one type of device may vary in the check-up items and maintenance period; however, the period may be determined through proper combination and adjustment according to the period planned for the entire system.

Table 1 Check-up Items for Preventive Maintenance

Item	Content
1. Check on signs of trouble	(1) Check on corrosion, cracks, rust, discoloration, and contamination (2) Investigation and measurement of abrasion, aging, and amount of wear (3) Check on loosening and leakage
2. Maintenance servicing on wearing parts	(1) Polishing of electric contact parts sliding parts (2) Lubrication on moving parts and friction parts (3) Cleaning of dusty, oily, or contaminated parts
3. Performance test	(1) Check on performance (2) Measurement and control of characteristic value

Maintenance of electronic devices

Priority is given to corrective maintenance. In this case, as previously mentioned $m < 1$ for most failures. While there are cases of $m > 1$ concerning individual parts, erroneous parts or errors in operating procedures are found in many cases; further, it is difficult from a practical viewpoint to execute characteristic control on each various part even in the case of $m > 1$. As a result, it is inevitable that importance be attached to corrective maintenance.
Since it is difficult to actually prevent trouble by maintenance, it proves effective to give the following consideration to devices:

1. The device should be equipped with a fail-safe function.

2. The device should preferably be equipped with a function for checking prior to start-up of operation.

3. Parts should be subjected to derating to a large extent.

4. Redundant system should be adopted (Preventive maintenance is effective in this case).

To perform adequate corrective maintenance. Regarding troubles in electronic apparatus, trouble-shooting is made in the field and defective modules are replaced with spare ones. These defective modules are generally repaired by the maker and returned to the maintenance department. It is necessary to take the following measures to minimize down time of the system caused by module trouble:

1. Detect and trace trouble precisely and rapidly.

2. Control and store an adequate quantity of spare parts and devices.

Spare parts and devices. It is advisable that spare parts and devices should adopt minimum units which permit determining a region by a technician of the maintenance department and thus facilitates replacement.
Regarding an adequate number of spare parts and devices to be stored, it is known that the quantity can be statistically determined from the relation between the expected trouble frequency and the out-of-stock ratio of necessary spare parts and device[5].

CONCLUSION

As a result of analyzing the field data on Hitachi's products, specifically devices for rolling stock and elevators, the Weibull parameter reveals various characteristics depending on the type of product and failure mode.

Relation between the Weibull parameter and maintenance method was studied and a difference in fundamental maintenance principles for mechanical, electric, and electronic apparatus was discussed.

The reader is reminded that the collection of field data was made by the manufacturer with the result that input data may be not necessarily sufficient in quality and quantity.

REFERENCES

1. Tsubouchi: On Control and Maintenance of Rolling Stock Based on Reliability Theory. Electric Railway 23-11
2. Ichida: Guide to Maintenance Engineering Nikkagiren
3. Regina C: "Optimal" Policy in a Maintenance Cost Problem. JORSA Vol. 15
4. W.M.Bassion: Increasing Hazard Function and Overhaul Policy. Annual Sym on R Vol.2, No.1
5. Omae, Morimura: Queuing Theory and Practice Nikkagiren

A STATISTICAL ANALYSIS OF MAINTENANCE SYSTEM SIMULATION*

INDEX SERIAL NUMBER – 1057

by

A.L. Goel, P.S. Nanda, and N. D'Sa

Department of Industrial Engineering and Operations Research; Systems and Information Science

Syracuse University
Syracuse, N.Y. 13210

Abstract

In this paper we present a statistical analysis and interpretation of the data from the simulation experiments conducted on a GPSS simulator developed for studying the unavailability and logistics support cost for ground electronics systems. The simulator considers three echelons of maintenance and provides for preventive maintenance of some components. Due to a large number of exogenous variables (both stochastic and deterministic) considered in this study, we have employed 2-level fractional factorial designs to achieve economy in experimentation and computer time. The approach taken is that of sequential experimentation, where fractions of full factorial designs are run sequentially while utilizing the results of analyses from the previous fractions.

1. Introduction

More than half of a system's total life-cycle cost can be attributed to its logistics costs, i.e. support, operation and training costs. The logistician is faced with the problem of making certain decisions, for example, the repair facilities required, the quantity and location of spares, the repair philosophy, maintenance personnel requirements, etc., in order to minimize his long term costs, subject to certain constraints such as reliability and availability. Thus, two major aspects of concern in the study of a system life-cycle are the maintenance cost and the system operational availability. If a penalty cost is charged against downtime or unavailability of the system, then cost becomes the main aspect of concern.

Recently, a number of studies have been aimed at determining the life-cycle costs of a maintenance system. The main variables considered have been spares, personnel, repair facilities required, level of repair and transportation. For a review of these studies see [4]. Most of these efforts do not take into consideration the fact that both the system unavailability cost and the maintenance cost are related to the various controllable and uncontrollable exogenous variables (e.g. mean time between failure, penalty cost per unit downtime etc.) of the maintenance system. Therefore, an optimal maintenance support plan should be based on a simultaneous study of the effects of these exogenous variables on the unavailability cost and the maintenance cost. In general, such an objective leads to a mathematical programming formulation of the problem, if the pertinent functional relationships are known. In general, however, these are hard to determine. This limitation, coupled with the fact that some of the exogenous variables are stochastic in nature, precludes an analytical solution of the above problem, and leads one to use computer simulation techniques.

In this paper, we present a statistical analysis and interpretation of the data from the maintenance simulation model developed in [3]. The model considers the exogenous variables alluded to previously, incorporates three echelons of maintenanance and is geared towards ground electronics systems. Although similar studies that consider one or more of the main exogenous variables have been reported [4], very little attempt has been made to determine the sensitivity of the endogenous variables to a range of values of the exogenous variables, or to perform a statistical analysis of their interrelationships.

In Section 2 we describe the two generic systems under investigation, viz. the hardware system and the maintenance system, while the specific system studied is discussed in Section 3. The role of designed experiments in simulation studies is illustrated by a 2^3 factorial design in Section 4. A series of 2^{7-4} Fractional Factorial are investigated in Section 5.

2. Description of the Generic Systems

A description of the components of both the generic hardware system and the generic system used for its maintenance is presented in this section.

The Generic Hardware System

The generic hardware system can be broken down as follows:

The system consists of a number of subsystems. Each subsystem is further composed of Higher Modular Assemblies (HMA's). The HMA's may be of two types, either compartmental (without modules) HMA's or modular (with modules) HMA's. Modular HMA's consist of many modules integrated into one HMA. Each module may also be of two types, either one that just requires alignment or one that comprises of many units requiring repair. Finally, the units consist of many printed circuits. The printed circuit is the smallest hardware unit in the system.

The Generic Maintenance System

This maintenance system consists of three echelons of maintenance viz., field, organization, and depot.

The corrective maintenance philosophy is assumed to be repair of all components except the printed circuits, which are discarded. The preventive maintenance philosophy is a block replacement policy subject to the constraint that spares are available in inventory.

A skeleton flow chart of the maintenance activities at each echelon is illustrated in Figures 1,2,3 and a description at each echelon is given below.

Corrective Maintenance at the Field Level

At this level, the logistics of the fault detection and correction operation proceed as shown in Figure 1. The field level being the most crucial of

*This research was supported by the Rome Air Development Center under contract No. F30602-71-C-0312.

the three echelons as far as downtime is concerned, a maximum allowable diagnosis time is allotted at this echelon. At the end of this period, the diagnosis process is stopped and the smallest possible faulty component (HMA or parent subsystem) is sent to either of the other two echelons of maintenance. When a fault in the system is detected (III)*, it is diagnosed to subsystem level. It has been assumed that this can be done within the allowable diagnosis time. Next, if the latter has expired, the faulty parent subsystem is removed and replaced by a spare if one exists (XIX and XX), and sent to the organization level for repair. If no spare exists, the system is down (XIX and XXII) until a spare arrives (XXIII), and a penalty cost is charged for the downtime. However, if time allows diagnosis at the HMA level, it is determined whether the HMA is compartmental or modular (VI). In either case, the faulty HMA is removed and replaced by a spare, if one exists (XI and XVI), and sent to another echelon for repair. If not, the parent subsystem is removed and replaced by a spare if one exists (X and XI). Again, downtime occurs if parent subsystem spares do not exist (XIII). The faulty compartmental HMA's are sent to the depot for repair (IX).

Corrective Maintenance at the Organization Level

At this level, the maximum allowable diagnosis time is greater than that at the field, because the level is not as crucial. The aim of the maintenance technicians is to diagnose faults to the module level, and to remove and replace faulty modules and send them to the depot for repair.

For subsystems sent here from the field (XXIV)**, if time does not permit diagnosis of the faulty HMA's, they are sent to the depot (XXV). If not, then it is determined if the HMA is compartmental or modular (XXVII). The former is removed and replaced by a spare, if one exists (XXVIII), and sent to the depot. The latter is diagnosed to the module level if time permits (XXXII and XXXIII) and removed and replaced by a spare, if one exists (XXXV), and sent to the depot. If time does not permit module diagnosis, then either the parent HMA or the parent subsystem is sent to the depot for further maintenance, depending on whether an HMA spare exists or not (XXXVII and XXXIX). We note that faulty compartmental HMA's are sent to this level at (B).

Corrective Maintenance at the Depot Level

This is the final maintenance echelon and there is no maximum allowable diagnosis time at this level. Compartmental HMA's sent here from the other levels are repaired (XLII)*** and returned to inventory, whereas modular HMA's are diagnosed and repaired down to the appropriate level (XLVI, XLVII and XLVIII). However, printed circuits are discarded and replaced. In each case, repaired components are returned to the appropriate inventory keeping in mind that the inventory at the field must be replenished before the inventory at the organization.

Preventive Maintenance

In this paper, preventive maintenance of only the compartmental HMA's at the field level is considered. A block replacement policy is used in which the compartmental HMA's are removed at times T, 2T, 3T,..., regardless of their failure history and replaced by spares from the inventory. If no spares exist, then the parent subsystem is removed and replaced by a spare. In either case, the replaced HMA is sent to the depot to bring it back to the "as good as new" condition.

System Descriptors

Because of the immense complexity of the maintenance system, a large number of variables and parameters need to be considered for a complete system description. Some of the exogenous variables considered in this paper are failure data, repair data, the configuration of the hardware system considered, repair philosophy, preventive maintenance time interval, etc. The status of the system is determined by looking at dynamic inventory levels, system downtime, etc. Some system parameters that need to be known are the mean and variance of failure and repair distributions, labor, transportation and unavailability cost per unit, transportation time between various echelons, etc. The endogenous variable considered is the total unavailability and logistic support cost.

3. The Specific System Studied

The configuration of the electronics system considered in this paper is shown in Figure 4. It has two types of subsystems and four types of HMA's. HMA (1) is a compartmental HMA. There are four types of modules and five types of units. Module (2) requires only alignment.

System Configuration Matrices

Associated with each component hierarchy of the system under study is a matrix that defines the number and type of components at that level of the hierarchy, the number and type of components at the next lower hierarchy, and the failure distributions for the components in this next lower level. The failure distributions in each case are defined by a function number. In the present case, there are three such matrices because there are three distinct hierarchies viz. (i) Subsystem-HMA, (ii) HMA-Module, (iii) Module-Unit. It is clear that, in general, a system with any number of hierarchies can be completely described by such matrices.

Variables and Parameters

Exogenous Variables. Seven exogenous variables are considered in this study. They are:

1. Mean time between failures.
2. Maximum allowable diagnosis time at the field level.
3. HMA diagnosis time.
4. Block replacement time interval.
5. Maximum allowable diagnosis time at the organization level.
6. Module diagnosis time per unit.
7. System down time penalty cost.

Endogenous Variable. The endogenous variable considered in this study is:

1. The unavailability and logistics support cost of the system.

Parameters. The parameter values used in this study are as follows:

1. Subsystem diagnosis time is two minutes for SS1 and three minutes for SS2.

---

*These numbers refer to Figure 1.
** These numbers refer to Figure 2.
*** These numbers refer to Figure 3.

2. Module diagnosis time, in minutes, is exponential with a mean equal to five times the number of units in the module.
3. Unit diagnosis time, in minutes, is exponential with a mean equal to three times the number of printed circuits in the Unit.
4. Time to remove and replace a subsystem or HMA is ten minutes.
5. Time to remove and replace a module is three minutes.
6. Time to diagnose the compartmental HMA is exponential with a mean of 20 minutes.
7. Time to align a module is exponential with a mean of 10 minutes.
8. The transportation time from the field level to the organization level is 30 minutes.
9. The transportation time from the organization level to the depot level is 120 minutes.
10. The transportation time from the field level to the depot level is 150 minutes.
11. The transportation cost is $0.20 per pound.
12. The average item weight is 50 pounds.
13. The labor costs are $9 per hour at the field and organization levels and $10 per hour at the depot level.
14. A factor of 4.3 is assumed to convert active labor hours to total labor hours.
15. The cost of having unit shortage is ten thousand dollars ($10,000).
16. The cost of storing a unit at the field level is $1000 and is $800 at the organization level.
17. There are 14 trucks, at a cost of $5 per truck per hour.
18. A total of 14 maintenance personnel are available, 8 at the field, 3 at the organization and 3 at the depot level.

4. A 2^3 Factorial Design and Analyses

The basic purpose of designing an experiment is to obtain the most information from the experimental data with least cost. Since computer simulation is indeed an experiment, careful consideration should be given to its design aspects. One of the aims of a simulation experiment is to study the system response over some region of operability in the factor space. To accomplish this objective efficiently with limited resources, a careful experimental design becomes crucial. Additionally, a good design provides desirable confounding patterns and computational ease.

In this section we illustrate the use of factorial designs by studying the effects of three independent variables on system unavailability and logistic support cost. Although a variety of designs can be considered, two level designs have proved very useful for initial investigations. Also, the results from such designs are easy to interpret. The three variables and the levels considered for each are given below:

For the sake of simplicity, these values are coded as follows:

$$x_1 = \frac{\text{MTBF} - 120{,}000}{60{,}000}$$

$$x_2 = \frac{\text{MADTF} - 120}{60}$$

$$x_3 = \frac{\text{PC} - 1.5}{1.0}$$

The experimental region delineated by these variables is shown geometrically in Figure 5. The eight points in this figure represent the $2^3 = 8$ runs necessary to consider all possible combinations of both levels of the three variables. The full factorial design along with two sets of simulated cost value, y_a and y_b, is given in Table 1. The replication provides an estimate of the error variance which is needed for evaluating the significance of the effects, as shown later in the paper.

The simulator was run for a total time of 2 million minutes, representing approximately 44 months of system operation. It should be pointed out that the manpower cost of $663.33 x 10^3 has been subtracted from all these values. This does not effect our analyses in any way, however.

The cost values associated with each point are shown geometrically in Figure 6. Thus, when MTBF= 180,000, MADTF= 180, and PC= 0.5, the cost values (in thousands of dollars) for the first and second simulations are respectively 121.9 and 100.0. The average cost is 111.0 as shown by point 4 in Figure 6.

Calculation of Main Effects

A geometric interpretation of the main effects is provided by the diagrams in Figure 7. Referring to Figure 7a, we note that when we increase the MTBF from 60,000 minutes to 180,000 minutes, while keeping the MADTF at 60 minutes and PC at 0.5 dollars per minute, i.e. when we move from point 1 to point 2, the cost decreases by 25.1 thousand dollars. In other words, when MADTF= 60 and PC= 0.5, the effect of increasing the MTBF from 60,000 to 180,000 is to decrease the cost by 25.1 units. Similarly, other changes in cost are obtained by subtracting the values at points 4, 6 and 8 from those at 3, 5 and 7 respectively. The average of these four differences is called the main effect of MTBF and is given by,

$$E_1 = \frac{1}{4}[(104.5-129.6) + (111.0-155.1) + (107.1-133.9) + (120.9-175.9)]$$

$$= -37.73.$$

VARIABLE		UNIT	LEVELS	
			LOW	HIGH
1. Mean Time Between Failures	(MTBF)	mins	60,000	180,000
2. Maximum Allowable Diagnosis Time at Field	(MADTF)	mins	60	180
3. Penalty Cost	(PC)	dollars/min	0.5	2.5

Alternatively, first the cost values along plane I and plane II are separately summed, the latter subtracted from the former, and the average taken to give us E_1. Using this method, we get

$$E_1 = \frac{1}{4}\,[(104.5+111.0+107.1+120.9) - (129.6+155.1+133.9+175.9)]$$

$$= -37.73.$$

This means that when the MTBF is increased from 60,000 to 180,000 within the experimental region shown in Figure 5, on the average the system unavailability and logistic support cost decreases by 37.73 thousand dollars.

Proceeding similarly, the main effects of MADTF and PC, obtained by considering planes (III, IV) and (V, VI) in Figure 7(b) and 7(c) respectively, are:

$$E_2 = \frac{1}{4}[(155.1+111.0+175.9+120.9) - (129.6+104.5+133.9+107.1)],\text{ and}$$

$$E_3 = \frac{1}{4}[(133.9+107.1+175.9+120.9) - (129.6+104.5+155.1+111.0)].$$

Calculation of Interaction Effects

A two factor interaction represents the effect on cost when two variables are changed simultaneously. The three 2-factor interactions in this case are: MTBF and MADTF, MTBF and PC, and MADTF and PC. A geometric interpretation of these effects is provided in Figure 8 by planes VII and VIII, IX and X, and XI and XII respectively. The interaction effect between MTBF and MADTF, for example, is obtained by taking the average of the difference in the sums of cost values on planes VII and VIII. Thus, we have

$$E_{12} = \frac{1}{4}[(129.6+111.0+133.9+120.9) - (104.5+155.1+107.1+175.9)]$$

$$= -11.8.$$

This value indicates that when MTBF is increased from 60,000 to 180,000 minutes and MADTF is increased from 60 to 180 minutes (the effect of PC is cancelled out.), the average change in cost is a decrease by 11.8 thousand dollars.

Proceeding similarly, the interaction effects between MTBF and PC (E_{13}) and between MADTF and PC(E_{23}) are -3.13 and 5.93 respectively.

Calculation of the Confidence Intervals for Effects

To ascertain the precision of the main and the interaction effects, a commonly used method is to calculate the appropriate confidence intervals. For a 2^3 design replicated twice, $100(1-\alpha)\%$ confidence interval for an effect is given by [2]:

$$E_i \pm t_{8,\alpha/2} \cdot s\ ,\ \text{where}$$

E_i is the calculated value of the effect,

s is the estimated standard deviation of the effect, and

$t_{8,\alpha/2}$ the appropriate value of the t-statistic.

The value of s is obtained from the replicated costs in Table 1 as shown in [2]. Using this formula, the 95% confidence limits for the main and the interaction effects are:

Effect	95% Confidence Interval
MTBF	(-73.925, -1.52)
MADTF	(-14.275, 58.125)
PC	(-26.825, 45.575)
MTBF-MADTF	(-47.975, 24.425)
MTBF-PC	(-39.325, 33.075)
MADTF-PC	(-30.275, 42.125)

These intervals imply that if different sets of observations are taken and the effects are calculated for each, then, 95% of these effects will lie in the appropriate intervals.

5. Fractional Factorial Designs and Analyses

As mentioned in Section 3, seven exogenous variables are of interest in this investigation. We consider two levels of each of these variables as given in Table 2. A full factorial design in these variables will require $2^7 = 128$ runs. Such a large number of runs is not only expensive, but is also unnecessary. Therefore, we employ carefully chosen fractional factorial designs to study the effects of the above variables.*

The first fraction is a 2^{7-4}_{III} design of 8 runs and is obtained from the following generators. Note that the variables are identified by their numbers as given in Table 2 rather than by the symbols.

$$I = 124,\quad I = 135,\quad I = 236,\quad I = 1237$$

These are called generators of the design because it is with these relations that we actually generate the design. For example, the generator I= 124, merely implies that 4= 12, i.e., the levels of variable 4 for each run are specified by multiplying the appropriate row elements in columns 1 and 2. (See Table 3). Hence 1.2= 4 or 1.2.4= (1.2).4= 4.4= I, where I is the column consisting of all elements equal to +1. Thus the standard form in which the generator is written in is I= 124.

The complete design, coded values for the variables and two replicates of simulated costs are given in Table 3. On analyzing the costs of Table 3, we obtain estimates of the combinations of main effects and higher order effects. Since these do not permit an easy interpretation, another design was obtained using the generators:

$$I = -124,\quad I = -135,\quad I = -236,\quad I = 1237$$

Here I= -124 implies that 4= -12, i.e., the levels of variable 4 for each run are specified by multiplying the appropriate row elements in columns 1 and 2 and reversing the signs(See Table 4).

The coded values and the replicated costs for this design are given in Table 4.

On combining the results from Tables 3 and 4, we get the following estimates:

* For details of this approach, the reader is referred to [1].

TABLE 1

DESIGN MATRIX, CODED VALUES AND THE COSTS FOR THE 2^3 FULL FACTORIAL DESIGN.

Run No.	Design Matrix			Coded Values			Cost ($x10^3$)		$\frac{y_a+y_b}{2}$
	MTBF	MADTF	PC	x_1	x_2	x_3	y_a	y_b	
1	60,000	60	0.5	-1	-1	-1	127.3	131.8	129.6
2	180.000	60	0.5	1	-1	-1	109.7	99.3	104.5
3	60,000	180	0.5	-1	1	-1	121.2	188.9	155.1
4	180,000	180	0.5	1	1	-1	121.9	100.0	111.0
5	60,000	60	2.5	-1	-1	1	131.1	136.6	133.9
6	180,000	60	2.5	1	-1	1	112.6	101.7	107.1
7	60,000	180	2.5	-1	1	1	127.1	224.6	175.9
8	180,000	180	2.5	1	1	1	137.1	104.8	120.9

TABLE 2

EXOGENOUS VARIABLES AND THEIR LEVELS

Variable Number	Variable Name	Unit	Level		Remarks
			Low	High	
1	Mean Time Between Failures (MTBF)	hrs	600	1000	Time between failures is exponential
2	Max. Allowable Diagnosis Time at the Field Level (MADTF)	mins	60	180	-
3	HMA Diagnosis Time (HMADT)	mins	25	50	Time is exponential
4	Block Replacement Time Interval (BRTI)	hrs	1200	1800	-
5	Max. Allowable Diagnosis Time at the organization level (MADTO)	mins	180	300	-
6	Module Diagnosis Time per unit (MDT)	mins	5	10	-
7	System Down-time PENALTY Cost (PC)	Dollars per min.	0.5	2.5	-

TABLE 3

DESIGN VALUES, CODED VALUES AND COST VALUES (FIRST FRACTION)

Run No.	Design Variables and Their Values							Codes Values							Cost ($\times 10^3$)		
	MTBF	MADTF	HMADT	BRTI	MADTO	MDT	PC	1	2	3	4=12	5=13	6=23	7=123	y_a	y_b	$\frac{y_a+y_b}{2}$
1	60,000	60	25	72,000	300	10	0.5	-	-	-	+	+	+	-	111.5	120.0	115.8
2	36,000	60	25	108,000	180	10	2.5	+	-	-	-	-	+	+	131.7	154.9	143.3
3	60,000	180	25	108,000	300	5	2.5	-	+	-	-	+	-	+	114.2	114.0	114.1
4	36,000	180	25	72,000	180	5	0.5	+	+	-	+	-	-	-	132.5	183.3	157.9
5	60,000	60	50	72,000	180	5	2.5	-	-	+	+	-	-	+	112.6	101.7	107.2
6	36,000	60	50	108,000	300	5	0.5	+	-	+	-	+	-	-	96.0	127.0	111.5
7	60,000	180	50	108,000	180	10	0.5	-	+	+	-	-	+	-	92.6	102.6	97.6
8	36,000	180	50	72,000	300	10	2.5	+	+	+	+	+	+	+	139.7	151.4	145.6

TABLE 4

CODED VALUES AND COSTS VALUES (SECOND FRACTION)

Run No.	Coded Values							Cost ($\times 10^3$)		
	1	2	3	4=-12	5=-13	6=-23	7=-123	Y_a	Y_b	$\frac{Y_a+Y_b}{2}$
1	-	-	-	-	-	-	-	105.0	94.2	99.6
2	+	-	-	+	+	-	+	115.8	96.1	106.0
3	-	+	-	+	-	+	+	114.2	103.4	108.8
4	+	+	-	-	+	+	-	124.7	129.2	127.0
5	-	-	+	-	+	+	+	101.4	111.6	106.5
6	+	-	+	+	-	+	-	127.3	131.8	129.6
7	-	+	+	+	+	-	-	104.7	114.2	109.5
8	+	+	+	-	-	-	+	133.5	125.0	129.3

$E_1 = 23.9$	$E_{12}+E_{37}+E_{56} = 8.6$
$E_2 = 8.8$	$E_{13}+E_{27}+E_{46} = -0.1$
$E_3 = -4.5$	$E_{14}+E_{36}+E_{57} = 0.6$
$E_4 = 6.5$	$E_{15}+E_{26}+E_{47} = -12.9$
$E_5 = -4.7$	$E_{16}+E_{25}+E_{34} = 5.4$
$E_6 = 4.9$	$E_{17}+E_{23}+E_{45} = -2.0$
$E_7 = 1.6$	$E_{24}+E_{35}+E_{67} = 7.0$

An estimate of the error variance was obtained from the two sets of replicated runs in Tables 3 and 4. On studying the above effects in the light of the error variance we find that effects E_1 and $E_{15}+E_{26}+E_{47}$ are the only ones which are statistically significant.

Since it is hard to tell which one of the two factor interactions is significant, it was decided to simulate another set of 8 runs using the generators:

I= -124, I= 135, I= 236, I= 1237.

Results from this run, coupled with those from the first two fractions, gave clear estimates of all 2-factor interactions involving variable 4. However, none of these was found to be significant. Therefore, the next fraction of 8 runs was simulated using the generators:

I= -124, I= 135, I= -236, I= -1237.

No conclusive evidence about the interaction effects was obtained from this fraction either. The next fraction was simulated by setting up a design from the generators:

I= 124, I= -135, I= 236, I= 1237.

Combining the results from the five fractions, the significant two factor interactions were $E_{35} = -9.9$ and $E_{67} = 14.3$.

Thus, from a total of 40 runs (not counting the replicates for the error variance estimate) we conclude that the three significant effects are those of MTBF and the 2-factor interactions between HMADT and MADTO and between MDT and PC.

It should be pointed out that the above results are valid only in the region of study as delineated in Table 2. Also, three and higher order interactions have been assumed to be non-significant in the above analyses. A similar approach can be used to study the effects of any set of exogenous variables in the desired ranges.

6. Conclusions

We have shown the methodology for setting up experiment designs and conducting statistical analyses for the simulation study of a ground electronics maintenance system. Although a specific system has been studied in this paper, a similar investigation can be conducted for any ground electronics maintenance system that fits the description of Section 2.

The main advantages of systematic experimentation are the economy in computer time, ease of statistical analyses and a clear interpretation of relationships between the exogenous and the endogenous variables.

References

1. Box, G.E.P. and J.S. Hunter, "The 2^{k-p} Fractional Factorial Designs", Technometrics, 3, Part I, 311-352; Part II, 449-458.

2. Goel, A.L., "Computer Based Statistical Modelling", unpublished notes, Dept. of Ind. Engg & O.R., Syracuse University, Syracuse, N.Y.

3. Nanda, P.S., A.L. Goel and N. D'Sa, "A GPSS Model for Ground Electronics Systems Maintenance", Proceedings of the 1972 Summer Computer Simulation Conference, June 14-16, 1972.

4. Paulson, R.M., R.B. Waina and L.H. Zacks, "Using Logistics Models in System Design and Early Support Planning", The Rand Corporation, R-550-PR, February 1971.

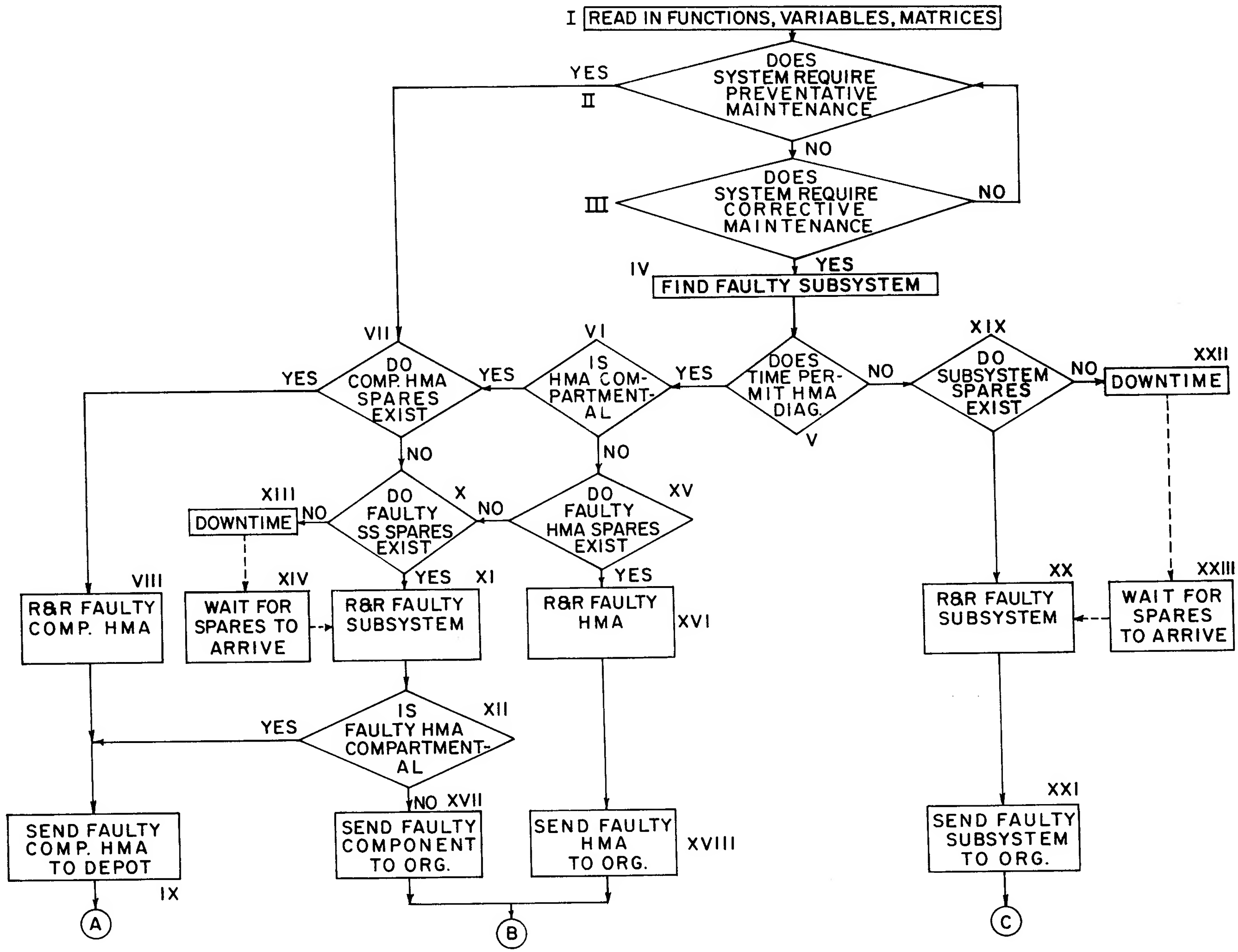


Figure 1. Corrective Maintenance at the Field Level

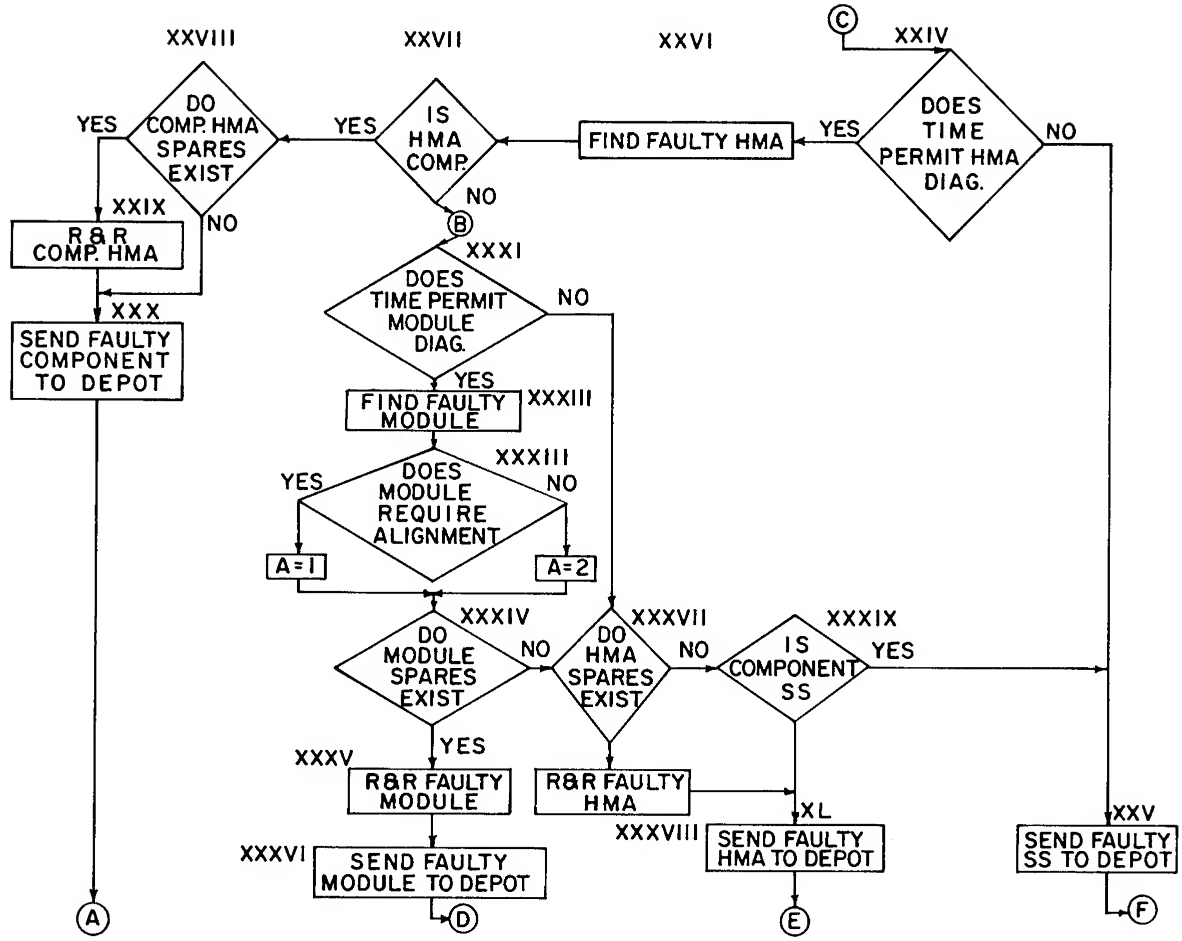


Figure 2. Corrective Maintenance at the Depot Level

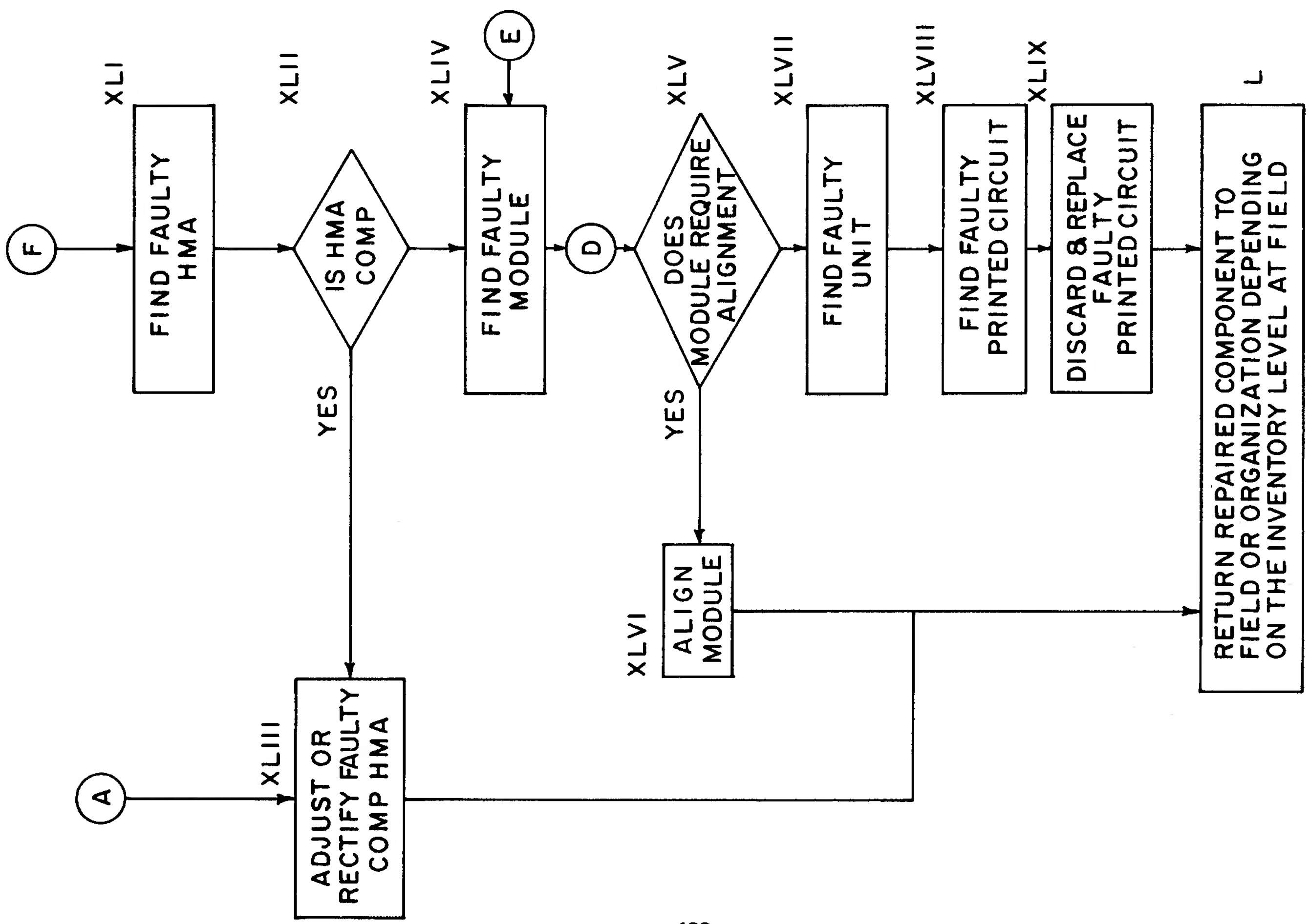
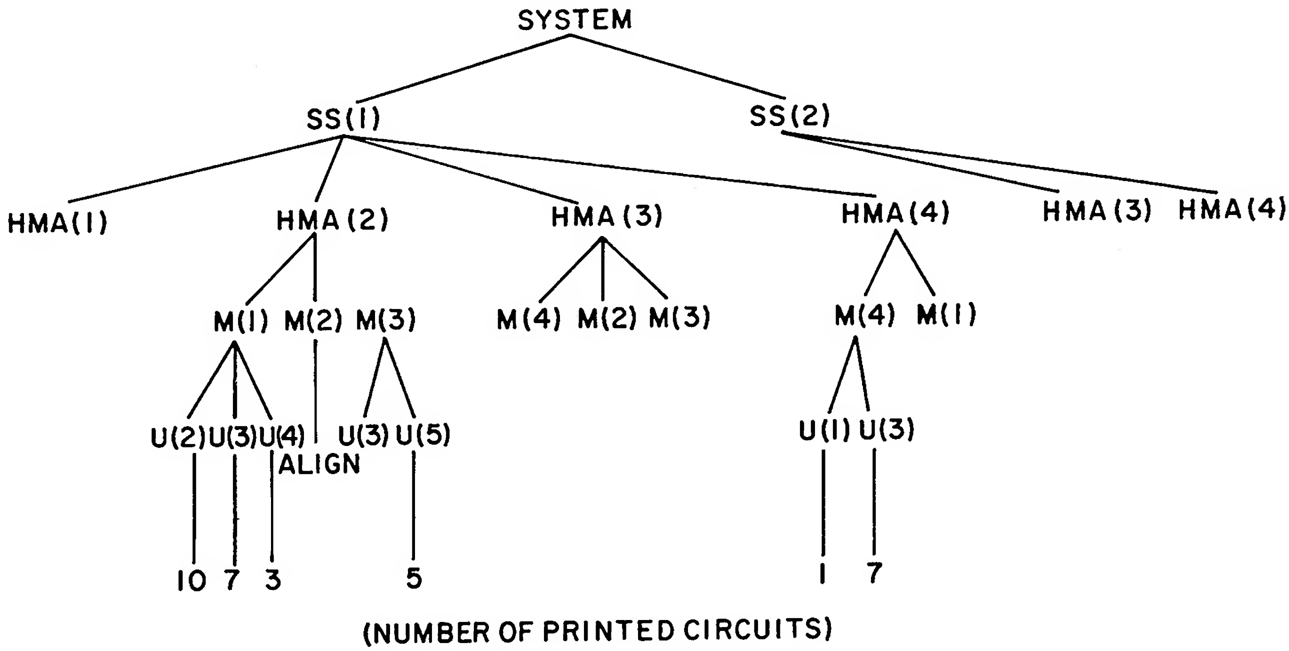
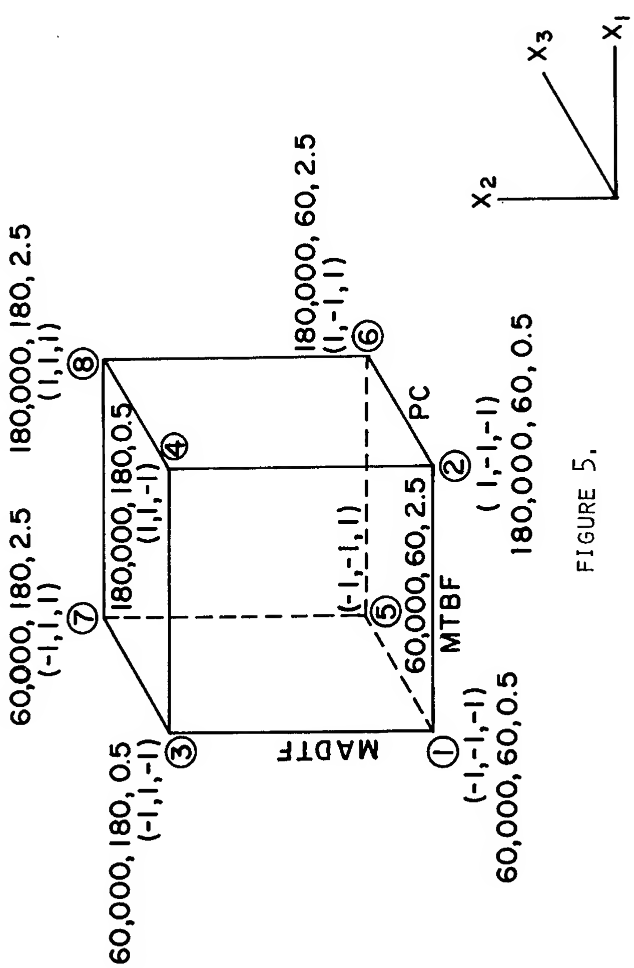


Figure 3. Corrective Maintenance at the Depot Level



SS(I) = SUBSYSTEM OF TYPE I, I = 1,2
HMA(J) = HMA OF TYPE J, J = 1,2,3,4
M(K) = MODULE OF TYPE K, K = 1,2,3,4
U(L) = UNIT OF TYPE L, L = 1,2,3,4,5

Figure 4. System Configuration for the System Study



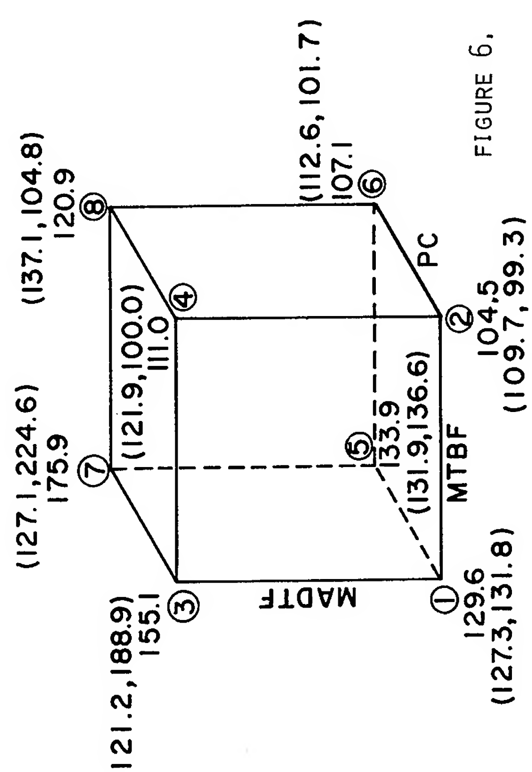


Figure 5. Geometrical Representation of Experimental Region

Figure 6. Replicated and Average Cost Values

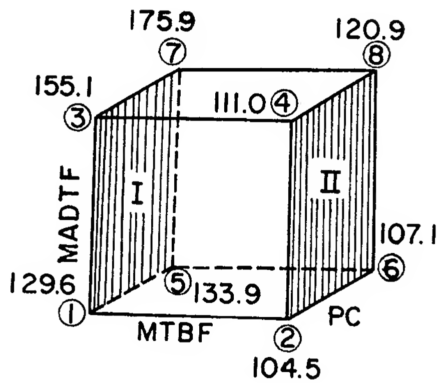


Figure 7. Geometric Representation of the Main Effects

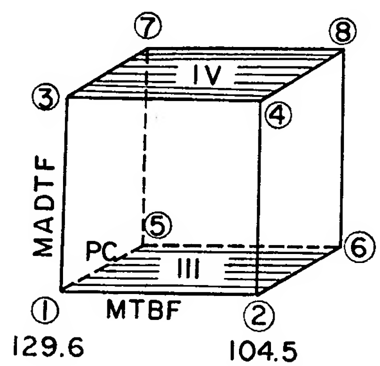


Figure 7A. Main Effect of MTBF

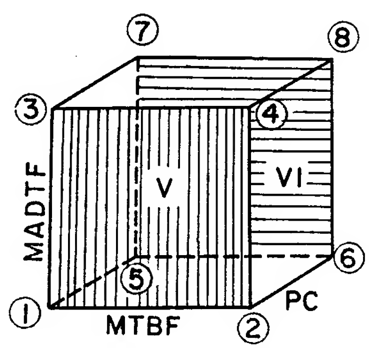


Figure 7B. Main Effects of MADTF

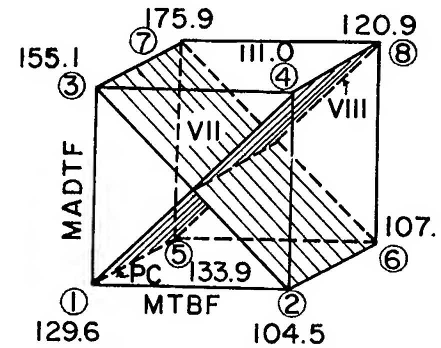


Figure 8. Geometric Representation of Interaction Effects

Figure 8A. Interaction Effects of MTBF and MADTF

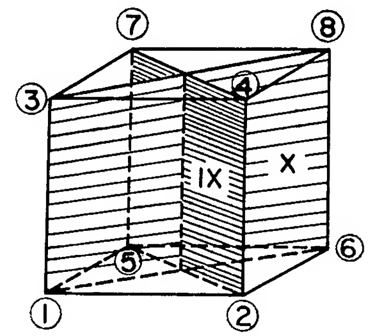
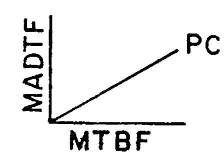


Figure 8B.

Interaction Effects of MTBF and PC



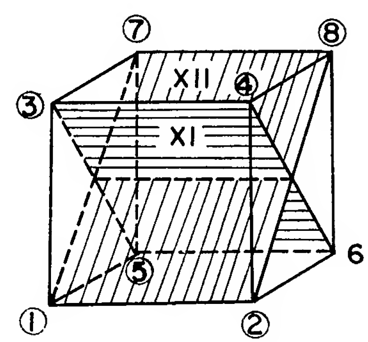


Figure 8C. Interaction Effects of MADTF and PC

QUALITY ASSURANCE FOR THE DATA PROCESSING INDUSTRY

Carl Sontz
TRACOR, Inc.
Rockville, Maryland 20850

INDEX SERIAL NUMBER — 1058

Introduction

Almost one percent of the total U. S. labor force, about 700,000 people, is working on tasks associated with the keying of data into computer systems.[1] Data preparation charges are as much as 40 to 50 percent of total EDP costs.

The assumption that computer numerical results are always accurate is incorrect.[2] Errors occur at many points in the processing cycle. Data preparation accounts for a large part of EDP costs, so that these errors have a major cost impact. In spite of the cost of errors, the data processing industry has been slow to adopt quality control techniques because its technical personnel and managers do not talk to their counterparts in the quality assurance sciences.

This paper discusses the impact of errors in data processing systems and gives a tutorial description of quality control techniques which will minimize the frequency of errors.

The following error sources will be discussed: incorrect recording of source data, improper coding, errors in data preparation, undetected program errors, hardware failures, undetected software (operating system and compiler) errors, incorrect or inefficient numerical computation routines, communication errors, power failures, operator errors, data base degradation.

Measures quantifying each source of error will be defined and the costs of reducing them will be discussed. The classic reliability measure, MTBF, is not appropriate for use in most cases and error rate is often more useful. Simple cost models for each error type will be presented.

The lack of adequate data and data sources seriously limits accurate cost analysis in this field. Several solutions to this problem are proposed.

Error Sources

Incorrect Recording of Source Data

Errors in original data may be impossible to correct because the original source may disappear. Frequently, this data is recorded in an office environment with no quality control. Coders and data input personnel must accept this data at face value, except for obvious errors such as incorrect format or decimal point. A great deal of time and money can be lost looking for errors in the EDP system which is, of course, the wrong place. It is assumed that source data error is not the responsibility of the EDP facility and no discussion will be given; however, the quality control techniques described here can be used to minimize this type of error.

Improper Coding

Coding is the assignment of alphanumeric codes to such items as categories, relationships between categories, classes of objects, etc. The codes are entered into the computer where they define data fields in which specific numeric information is stored. For example, a drug distribution firm can record data on thousands of drugs defined in the computer by alphanumeric codes rather than by name. The use of alphanumeric codes simplifies information retrieval, especially in older tape-oriented business systems. Coding errors are quite significant and error rates of 3 to 5 percent are common.[3]

The coding error rate, $r(CD)$, can best be measured by the ratio of items incorrectly coded, $N_e(CD)$, to total items coded, $N(CD)$, in some convenient time period. A measure of operator productivity, called throughput, is the number of items coded per unit time by that operator. For the i'th operator, this is $N_{oi}(CD)$. A measure of operator effectiveness is the ratio of the average error rate to the throughput for that operator. This measure permits management to consider both error rate and throughput in evaluating a coder's efficieny.

Verification is required in a coding operation. The verification error rate, $r_v(CD)$, is defined as the ratio of items incorrectly verified to total items verified. Two types of verification errors are possible: (1) incorrect items are diagnosed as correct, and (2) correct items are diagnosed as incorrect. Thus,

$$r_v(CD) = \frac{\text{Type 1 errors + Type 2 errors}}{\text{Total items verified}} \quad . \quad (1)$$

A measure of verification productivity is the number of items verified per unit time, $N_v(CD)$. A measure of verification effectiveness is the ratio of verification error rate to verification throughput.

The costs of coding can be estimated from the following general relationship:

$$C(CD) = C_o(CD) + C_{ot}(CD) + C_s(CD) + C_v(CD) + C_{vt}(CD) + \Delta[C_o(CD) + C_s(CD) + C_v(CD)] + C_q(CD) \quad (2)$$

where

$C_o(CD)$ = the cost of operators
$C_{ot}(CD)$ = the cost of training operators
$C_s(CD)$ = the cost of supplies
$C_v(CD)$ = the cost of verifiers
$C_{vt}(CD)$ = the cost of training verifiers
Δ = an operator which indicates the extra costs arising from error correction
$C_q(CD)$ = the cost of quality assurance provisions, including special quality training, re-training, monitoring.

Errors in Input Preparation

The largest source of error in data processing system lies in input preparation. Operators key data from source documents into a computer readable form. The most common form of data entry is keypunching, which includes keying, verification, computer checks of data, and data re-entry.[4]

Verification is performed by a verification operator who keys the data a second time. If the keyed data items differ, a decision is made as to whether the original data entry or the verification was incorrect. Items to be re-entered are sent back to the data entry station and the process is repeated.

The error rate for input preparation, $r(IP)$, is the ratio of incorrect characters keyed to total characters keyed. Three error rates are of interest: $r_{oi}(IP)$, the error rate of the i'th operator; $r_{vi}(IP)$, the error rate of the i'th verifier; and $r_s(IP)$, the error rate for data which finds its way into the computer system.

Another useful parameter is throughput, $N(IP)$, or the number of key strokes per unit time. Three throughputs are of interest: $N_{oi}(IP)$, the throughput of the i'th operator; $N_{vi}(IP)$, the throughput of the i'th verifier; and $N_s(IP)$, the system throughput.

A measure of efficiency of operation, $E(IP)$, is the error rate per unit throughput. Again, three efficiencies can be derived: $E_{oi}(IP)$, the efficiency of the i'th operator; $E_{vi}(IP)$, the efficiency of the i'th verifier; and $E_s(IP)$, the efficiency of the system.

A general equation for input preparation costs for any keying operation is:

$$C(IP) = C_o(IP) + C_{ot}(IP) + C_e(IP) + C_v(IP) + C_{vt}(IP) + C_s(IP) + \Delta[C_o(IP) + C_e(IP) + C_v(IP) + C_s(IP)] + C_q(IP) \quad (3)$$

where

$C_e(IP)$ = the cost of input preparation and equipment

and the other cost terms have been previously defined for coding.

The number of keying operators, KO, required on a one-shift operation can be estimated from:[5]

$$KO = \frac{\text{Monthly Data Volume}}{\text{Operator Monthly Production}} = \frac{[1 + V + r(IP)]\,A}{U \times D \times S \times F} \quad (4)$$

where

$r(IP)$ = the input preparation error rate
V = the verification factor
A = the monthly volume of source data in characters
U = the useful hours/day/average operator (typically, U is 6)
D = the number of working days per month (typically, D is 20)
S = the average keying operator speed in strokes/hour (typically, S is 6500)
F = an equipment speed factor relative to keypunch speed (F is 1 for keypunch and 1.3 for key-to-tape and key-to-disc).

The verification factor, V, is given by:[5]

$$V = (PK \times KS) + (PV \times VS) \quad (5)$$

where

PK = the percentage of input data which is key verified (typically, PK is 1 for keypunch and less than 1 for key-to-disc, etc.)
KS = the key verification speed factor (KS is a constant factor of 1)
PV = the percentage of input data which is visually verified (typically, PV is 0 for keypunch and greater than 0 for key-to-disc, etc.)
VS = the visual verification speed factor (typically, VS is 0.3, relative to keying).

V is 1 for keypunch and less than 1 for key-to-tape, key-to-disc, etc. V=1 means that all input data is verified.

In a keypunch operation, typical first re-entry costs may run as high as 8 percent of initial entry costs for an error rate of 3.6 percent.[4] Second re-entry costs may be as much as 24% of initial entry costs for a 0.4% error rate because of very high clerical research costs.

The cost of correcting data input errors after they reach the computer can be as much as $10 per error.[4]

Undetected Program Errors (Application Programs and Software)

Undetected program errors, whether occuring in application programs or software, which pass through program production are hard to find and can cause considerable difficulty before being detected.

The error rate for program errors, $r(P)$, is defined as the ratio of incorrect statements, $N_e(P)$, to the total number of statements, $N(P)$. Program reliability can be defined as:

$$R(P) = 1 - \frac{N_e(P)}{N(P)} \quad . \tag{6}$$

Reliability can be used as a measure of the efficiency of the producer and to compare the average effects of quality control procedures. However, for cost analysis, the important characteristic is the total number of errors because a program should be error-free. In a very large system, software errors in loops which are infrequently used may not be uncovered for a long time and may have very little impact. The goal, however, is still zero errors.

If a constant number of errors are corrected on each correction pass, the number of passes required to remove all errors, $P_c(P)$, is equal to the ratio of the number of errors, $N_e(P)$, to the number of errors corrected per pass, $N_c(P)$.

For a program with $N_e(P)$ errors, the cost of errors is:

$$C_e(P) = \sum^{N_e(P)} \left[C_{eo}(P) + C_{ec}(P)\right] \tag{7}$$

where

$C_{eo}(P)$ = the cost of occurrence of the error
$C_{ec}(P)$ = the cost of correcting the error.

The term $C_{eo}(P)$ represents the sum of wasted manpower, lost equipment time, law suits, and ill will. The term $C_{ec}(P)$ represents the sum of manpower, equipment, and supply costs required for error correction.

Equipment Failures

Equipment failures cause errors in the data processing outputs. An equipment failure occurs when a hardware failure results in a data processing error. An equipment incident occurs when a hardware failure does not result in a data processing error.

Equipment failure can be characterized by a failure rate, λ, in failures per unit time.[6] The reliability of a single equipment can be computed directly from the simple exponential reliability law. The reliability of complex computer systems can be computed, depending on the system configuration, using well-known reliability prediction techniques.[6]

A significant parameter is the mean time between failures, MTBF, which can be computed as a function of the system reliability.

The cost of equipment failures consists of the cost of lost time, incorrect results, and corrective maintenance. If the average repair time per failure is $\bar{\mu}$, the average time spent in repair during some calendar period T is approximately

$$T_r = \lambda\bar{\mu}T \quad . \tag{8}$$

The cost of corrective maintenance is

$$C_{cm}(E) = \sum^{f} T_r \, C_{cm/t}(E) \tag{9}$$

where

f = the number of failures
$C_{cm/t}(E)$ = the cost per unit time for corrective maintenance.

The cost of preventive maintenance is usually included in the equipment lease cost.

Incorrect or Inefficient Numerical Computation Routines

Computer programs are often used to process numerical data, make calculations, and produce numerical results which can be used in planning future operations. The quality of the output depends on factors such as the manner in which series summations are truncated, the types of numerical approximation algorithms used, etc.

A variety of computational errors can occur, including truncation errors, rounding errors, subtraction errors, errors in periodic functions, shifting loss, and numerical integration errors. These errors can produce a variety of output "errors" ranging from wrong answers and insignificant figures in the final answer to incorrect variances in analyzed statistical data. There is a substantial literature on this subject and it will not be discussed further.[7]

Communication Errors

Communication errors are unavoidable.[8] Electrical noise results in a residual error rate which, although small, can never be eliminated. The communication error rate, $r(C)$, is a function of the channel parameters, sig-

nal-to-noise ratio, and the error detection and correction techniques used.

Communication errors are very infrequent. Unless an obvious noise condition develops on the communications lines, they may not be distinguishable from other errors. Except for obvious factors, such as the cost of special equipment for improving signal-to-noise ratio, it is very difficult to pin down exact cost factors.

Communication errors fall into two general classes: systematic and random errors.[8] Systematic distortions, such as amplitude attenuation, delay distortion, frequency offset, and bias distortion, are characteristic of the transmission system and are of a relatively constant nature. Therefore, they can often be compensated for.

Random errors are caused by white noise and impulsive noise. They can be predicted only on a probabilistic basis. They are difficult to compensate for and require the use of error detection and correction systems.

White noise results from the thermal exitation of the electrons in a communication system. It is unavoidable and a part of all communications systems.

Impulse noise is a major source of transmission errors. Noise amplitudes may be quite high and of sufficient duration to affect several characteristics in a transmission. Impulse noise originates from external and internal sources such as switching actions in telephone offices through which the circuit is routed, ringing signals and test tones on adjacent circuits, lightning, intermodulation products, cross-talk, poor or dirty contacts in the transmission equipment, or transmission echoes.

Bell System technical data indicates an average of three interruptions per circuit per year, averaging 1.8 hours each.[8] Of all interruptions, 80% are of less than 2 hours duration. The Bell System has made three major studies of error rates in switched data transmission since 1960[9,10,11]. Based on this data, long term average error rate for private line voice channel can be taken as 1 in 10^5 or better during normal transmission conditions.

Power Failure

Power failure is becoming an increasingly important consideration in data processing systems.[12] The great demand for electrical power has resulted in frequent power failures and fluctuations during peak load periods, especially in the major metropolitan areas.

Short term power fluctuations, such as spikes, dips, and flicker, of greater than a millisecond duration which are not noticeable through dimming lights or other indications can still garble data which is being transmitted in a computer or through I/O. Errors caused by power fluctuations may remain undetected until after a program has been run and the results have been put to use. Even if the error is located, a great deal of time and money may be required to correct it.

Short term power fluctuations may occur several times per day. There is little data published at this time which can be used to estimate frequency of occurrence and recovery time and costs.

Power outages may occur a few times a year, knocking out a computer installation for hours or even days, depending on the nature of the occurrence.

The important characteristics for measuring the effects of complete outages are rate of occurrence, $r_o(PF)$, in outages per year, and average time per occurrence, $t_o(PF)$, in hours. The total expected outage time is

$$T_o(PF) = r_o(PF)\ t_o(PF) \ . \qquad (10)$$

If the cost per outage is $c_o(PF)$, the expected total outage cost per year is

$$C_o(PF) = r_o(PF)\ c_o(PF) \ . \qquad (11)$$

The characteristics for measuring the effects of power fluctuations are also rate of occurrence, $r_f(PF)$, and average time per occurrence, $t_f(PF)$. If the cost per fluctuation is $c_f(PF)$, the expected total cost per year of short term power fluctuations, $C_f(PF)$, can also be computed.

The total cost of power failure is the sum of the cost of fluctuation and outages:

$$C(PF) = C_o(PF) + C_f(PF) \ . \qquad (12)$$

The cost of these outages must be balanced against the cost of providing back-up power.

Operator Errors[13]

Operator errors can be measured by an error rate, $r(O)$, measured in numer of errors, $N_e(O)$, per unit time, T. The time period, T, would normally be one hour, but any other length can be used.

For the i'th operator who settles down to an average error rate of $r_i(O)$, the expected number of errors in time period T is

$$N_{ei}(O) = r_i(O)\ T \ . \qquad (13)$$

The total cost of operator errors is:

$$C(O) = N_e(O)\ c(O) \qquad (14)$$

where

$c(0)$ = the cost per operator error.

The cost per operator error includes factors such as the cost of re-running the program. The exact combination of factors which determines costs must be determined separately for each installation.

The effect of operator errors on the user of a system is different for batch and real-time systems. In a batch system, the computer operator can simply re-run the program. The cost to the computer facility is the cost of lost time, including extra computer time. If the customer receives his output on time, the computer facility receives no further penalty.

In a real-time system, the effects of operator error are more serious and immediate. The user feels the effect of an error at once and begins to lose money instantly when a failure occurs. The overall cost to the computer installation can be great. Bad will and lost customers may cost a lot more than lost time.

Data Base Degradation[14]

Data base degradation is a subject which has received small consideration in spite of its importance. Errors can be introduced into a good data base whenever it is interrogated. The data base can be altered by bugs in application programs, transmission errors, terminal errors, equipment errors during read/write operations, etc.

A measure of data base erosion, $r_e(DB)$, is the ratio of the number of characters in the data base which are changed in a unit time, $N_e(DB)$, to the total number of characters in the data base, $N(DB)$.

The effects of data base degradation may not be obvious. They may result in a variety of ills up to and including incorrect management decisions.

The cost of data base degradation is very difficult to measure directly. It may not be possible to relate the results of all errors to the errors that caused them. The cost of restoring the data base is a function of manpower, time, computer time costs, and overhead.

If the original data is disposed of, data base restoration is impossible. Therefore, a duplicate of all critical data base fields should be kept.

Error Management and Quality Control

General

There are a large number of error sources in a data processing system which can lead to a substantial build-up in errors. A quality assurance program should be developed in every data processing organization to minimize errors.

An important concept in quality control is the acceptable quality level (AQL), which, for a data processing system, is the maximum error rate that can be tolerated. A simple quality control technique using an AQL derived from historical data is the fraction-defective or p-chart technique. The application of this technique to data processing is outlined below.[15]

First, an AQL for error rates is developed from historical data. Then, an upper control limit (UCL) is computed from

$$UCL = AQL + N_{(1-\alpha)}\sqrt{\frac{AQL(1 - AQL)}{n}} \qquad (15)$$

where

$N_{(1-\alpha)}$ = the $(1-\alpha)$ cut-off point of the standard normal distribution

α is usually taken to be 0.1, 0.05, or 0.01

n = the sample size of incoming data.

The error rates computed for consecutive samples of a given size are then plotted. A sample is accepted if its error rate is less than the UCL. Increasing trends in error rate may indicate a deteriorating data quality. More detailed quality control procedures will be described below.

Quality Assurance for Computer Programs and Software

A number of quality control techniques have been developed to reduce program and software errors.[17] Configuration control procedures, debugging techniques, and the use of base line input data for test runs are all useful. The user of a program or software should develop quality checks whose objective is to reduce errors to zero.

Quality Assurance for Power Failures

Errors due to power failure or power fluctuations can be reduced to almost zero by means of on-site alternate power systems, voltage regulation devices, and interference reduction circuitry.[12] These techniques are expensive but provide excellent protection. Ordinary quality control procedures are of little value in combatting power failures.

Quality Assurance for Communication Errors

Communication errors can be reduced by installing error detection and correction systems of two types: forward-acting and re-transmission.[8]

Forward-acting systems use parity bits added at the transmitter which permit error correction and detection at the receiver. The receiver is designed to extract the original information signal and detect and correct errors.

Re-transmission systems also use parity bits. When an error is detected by the receiver, a signal is sent to the transmitter and the block of data is re-transmitted. Fewer parity bits are required in re-transmission systems because the receiver does not require the extra bits for error correction. The re-transmission system is the one most commonly used commercially.

Modern error detection and correction codes are quite effective. Starting with a channel error rate of 1 in 10^5, residual error rates of as little as 1 in 10^7 to 1 in 10^9 can be obtained.

Channel loss can cause a large block of data to be lost. This problem can be reduced by using redundant transmission paths.

Quality Assurance for Data Base Degradation

Several techniques are available for minimizing data base degradation, namely: keeping a duplicate copy of the data base, copying the entire data base into some permanent record form at the end of some time interval, and audit trails.[14]

The first solution is to keep a duplicate data base and to update both simultaneously. This is very expensive for large data bases, and, unless much of the computer system is duplicated, the full protective value of the duplicate data base may not be obtained. Power failures and other failures in non-redundant parts of the system can wipe out both data bases.

In the second approach, known as "disc dump," the entire base is copied onto a medium such as punched cards. An hour or more may be required for copying a large data base. Another drawback is that the system is unprotected between dumps. Also, repeated dumping and loading may deteriorate the data base. If diagnostic capabilities are built into the disc dump program, the procedure can be effective. These diagnostics should be designed to check that the file structure is undamaged.

In the audit trail approach, the system is designed to keep some combination of the following: the text of terminal input messages; copies of data base records before they are updated; copies of data base records after they are updated. This technique can also be used to prevent data base degradation.

Quality Assurance for Source Data Preparation, Coding, Keying Operations, and Machine Operators

General. Quality assurance procedures have been developed which can be applied directly to data processing operations involving human actions, such as source data preparation, coding, keypunching, typing, key-to-tape, key-to-disc, editing, and the activities of machine operators.[18,19]

The procedures are designed to improve quality and reduce the number of corrections. This approach has proven to be less expensive than those which rely exclusively on acceptance sampling techniques.[18]

All operators must be trained on the task they are to perform. After the training period, the operators must be qualified. During qualification, the work lots produced are monitored and accepted or rejected as if they were production lots. At the same time, one of two rules described below can be used to judge the operator's ability. Personnel who fail to meet the qualification requirements can be re-trained or dismissed. Those who pass the qualification tests are then placed into the system. Process control procedures are then used to monitor their performance.

Qualification. During the qualification period, lots can be accepted or rejected for use as regular production items. A random sample should be selected from each work lot produced by a given operator and a decision made to accept or to reject it based on the number of errors in the sample. Rejected lots are completely reworked and defective items are corrected.

As decisions are made on lot quality, one of the following two rules is used to qualify the operator:[18]

(1) Qualify if s successive work lots are accepted within a maximum of d inspected.

(2) Qualify if f or fewer work lots are rejected in d inspected.

Operators who do not qualify can be dismissed or re-trained. The re-trained operators can be given a second chance to qualify.

Accepting or Rejecting Lots. Several assumptions are made about production and inspection, namely:[18]

(1) A series of successive work lots is produced by a continuing process.

(2) The process to be evaluated is the sequence of work lots produced by an individual operator.

(3) The work lots produced are expected to be of the same quality with a true fraction defective, P.

(4) The sample is small relative to the size of the lot and the binomial distribution can be used to compute the probability of acceptance in a single sample.

(5) The inspection is without error. (Models which include inspector errors are discussed later.)

With these assumptions, the probability of accepting a lot after evaluating a single sample of n items is:

$$L_p = \sum_{x=0}^{c} \binom{n}{x} P^x (1 - P)^{n-x} \tag{16}$$

where

c = the maximum allowable number of defectives
P = the true fraction defective.

More detailed information on sampling plans can be found in 23 and 24.

The First Qualification Rule. The first qualification rule is that s successive work lots must be accepted within a maximum of d inspected on a sample basis. The probability of s successive accepted work lots within a maximum of d inspected is given by:[18]

$$Q_{d,s} = L_p^s + \sum_{i=1}^{r} M_i (L_p)^{is} (1 - L_p)^i$$
$$- \sum_{i=1}^{r-1} (N_i)(L_p)^{(i+1)s} (1 - L_p)^i \tag{17}$$

where

r = the largest integer in the quotient d/s

$$M_i = (-1)^{i+1} \binom{d - si}{i} \tag{18}$$

$$N_i = (-1)^{i+1} \binom{d - (i+1)s}{i} . \tag{19}$$

The probability $Q_{d,s}$ is shown in Table 1 for several values of L_p, d, and s.

The Second Qualification Rule. The second qualification rule is that an operator is qualified if f or fewer work lots fail the sample inspection within d lots inspected. The probability of f or fewer rejected work lots within d inspected is given by the binomial formula:[18]

$$S_{d,f} = \sum_{x=0}^{f} \binom{d}{X} L_p^{d-X} (1 - L_p)^X \tag{20}$$

where

X = the number of work units rejected in sample inspection (X = 0,1,2,...f).

Probabilities of qualifying $S_{d,f}$ for various values of L_p, d, and f are given in Table 2.

Rules for Monitoring the Performance of Qualified Producers. After workers are qualified, a quality control plan to control their performance must be used. When a worker begins qualified production, he receives a bonus of C points in his "account." The selection of C is based on the same factors which influence the choice of a qualification rule. A point is added to the worker's account whenever a decision is made to accept a work lot. The acceptance rule is that c or fewer defective items be found in the sample of n. A point is deducted when c+1 or more defective items are observed. The formula for the probability of surviving D decisions , starting with C points, is:[18]

$$U_{D,C} = \sum_{x=F}^{D} \left[\binom{D}{\frac{D-x}{2}} - \binom{D}{\frac{D-x}{2} - C} \right] L_p^{\frac{D+x}{2}} (1-L_p)^{\frac{D-x}{2}} \tag{21}$$

where

$$F = D - 2 \ (D + C - 1)/2 \tag{22}$$

is the lower limit. The value x denotes the net score which is cumulated over the range F to D, inclusive. The value of $U_{D,C}$ for various values of L_p, D, and C is given in Table 3.

Two more properties of the decision rule used to monitor the efficiency of personnel used in production are the number of credits expected to be accumulated by a worker remaining in the process at the end of D decisions, and the expected duration of the process given that a worker is removed on or before the D'th decision made on his work.

Expected Points If the Worker Survives. An equation for computing the average number of points a worker collects over D decisions, given that he survives, is given by:[19]

$$E(C_{D,C}) = \frac{1}{U_{D,C}} \left\{ \sum_{x=F}^{D} x \left[\binom{D}{\frac{D-x}{2}} - \binom{D}{\frac{D-x}{2} - C} \right] L_p^{(D+x)/2} (1-L_p)^{(D-x)/2} \right\} + C . \tag{23}$$

This equation can be simplified so that a binomial table or computer can be used to compute the expected number of points. Values for $E(C_{D,C})$ are given in Table 4 for various values of L_p, D, and C. The term "$C_{D,C}$" is used in this table instead of $E(C_{D,C})$.[19]

The expected process length (in numbers of decisions), given that a worker is removed on or before the D'th decision on his work, has been computed. Values of $Z_{D,C}$ for various values of L_p, D, and C are given in Table 5.

The Effect of Inspection and Correction Error. In the discussion above, it was assumed that inspection and correction were perfect. Of course, this is impossible. Several relationships for taking inspection and correction error into account are described below.[20,21,22]

With no error in inspection, the binomial equation (16) can be used to compute the probability of accepting a work lot of n items. This formula must be modified to account for inspection and correction error.

The effective fraction defective, P', is assumed to be a linear function of the true fraction defective, P:

$$P' = (1 - \beta_1)P + \alpha_1(1 - P) \tag{24}$$

where

α_1 = the probability of classifying a non-defective item as defective

β_1 = the probability of classifying a defective item as non-defective.

If P' is substituted for P in (16), the following is obtained:

$$L_{p'} = \sum_{x=0}^{c} \binom{n}{x} \{\alpha_1 + (1 - \alpha_1 - \beta_1)P\}^x \{\beta_1 + (1 - \alpha_1 - \beta_1)(1 - P)\}^{n-x} . \tag{25}$$

α_1 and β_1 represent the probability of making Type 1 and Type 2 errors, respectively.

The probability of classifying a non-defective item as a defective can often be ignored. If α_1 is zero, the probability of correctly classifying an item in the sample,

$$\Pi = 1 - \alpha_1 - \beta_1 \tag{26}$$

becomes

$$\Pi = 1 - \beta_1 . \tag{27}$$

When Π is substituted into (25), the formula becomes:

$$L_{p'} = \sum_{x=0}^{c} \binom{n}{x} (\Pi P)^x (1 - \Pi P)^{n-x} . \tag{28}$$

If $\Pi = 1$, corresponding to correct classification, (28) reduces to (16). The ability to discriminate between good and poor quality work is reduced by verification errors. For a given fraction defective, the probability of accepting the work lot is increased.

Average Outgoing Quality. The average outgoing quality, AOQ, is another measure of the quality of the final work product of a data processing installation. The AOQ is affected by inspection and correction error. Formulas which include the effect of inspection and correction error are given below.[22]

The effective AOQ is:

$$AOQ = \frac{M(N-n)PL_{p'} + Mn\beta_1 P + M(N-n)\beta_2 P(1-L_{p'})}{MN}$$

$$= (\frac{N-n}{N})PL_{p'} + \frac{n}{N}\beta_1 P + (\frac{N-n}{N})\beta_2 P(1-L_{p'}) \tag{29}$$

where

M = the number of work units processed

N = the number of items per work unit

n = the number of items sampled from the work unit

β_1 = the probability of classifying a defective item incorrectly in sample inspection $(0 \leq \beta_1 \leq 1)$

β_2 = the probability of classifying a defective item incorrectly when correcting items in a rejected work unit which is reinspected $(0 \leq \beta_2 \leq 1)$

P = the true fraction defective in a work unit

$M(N-n)\ _2P(1-L_{p'})$ = the expected number of defective items remaining in rejected work units after the correction process

$Mn\ _1P$ = the expected number of defective items remaining in the sample of items selected from work units

$M(N-n)PL_{p'}$ = the expected number of defective items remaining in accepted work units after sampling inspection is completed.

(29) can be simplified for various conditions. An AOQ equation is given below for several conditions.

In the case where no error is made in classifying and in correction, β_1 and β_2 are assumed to be zero and $L_{p'}$ equal to L_p. This is perfect inspection. The AOQ is given by:

$$AOQ = (\frac{N-n}{N})\ PL_{p'} . \tag{30}$$

In the case where there is a β_1 error in classifying and zero error in correction, β_1 is a fraction between zero and one and β_2 is zero. The AOQ is given by:

$$AOQ = (\frac{N-n}{N}) PL_{p'} + \frac{n}{N} \beta_1 P . \quad (31)$$

In the case where there is a β_1 error in classifying and β_2 correction error, both β_1 and β_2 are fractions between zero and one. Under these conditions, (29) can be used. If β_1 equals β_2, the AOQ is:

$$AOQ = (\frac{N-n}{N}) PL_{p'} + \frac{n}{N}\beta_1 P + (\frac{N-n}{N})\beta_1 P(1-L_{p'}) . \quad (32)$$

If $\frac{P'}{1-\beta_1}$ is substitued for P, then

$$AOQ = (\frac{N-n}{N}) P'L_{p'} + \frac{1}{1-\beta_1} \; P' . \quad (33)$$

Availability of Data

Although many workers and managers in the data processing industry recognize the high cost of errors, very little published data is available. Error data and the cost of errors are treated as private information which cannot be revealed to competitors or customers. Much of the available data is published by manufacturers who wish to prove the advantages of specific brands of hardware.

The data processing industry must develop professional personnel skilled in applying quality assurance techniques to their industry. One of their first tasks is the development of statistically valid data on errors and the cost of errors.

A very useful project would be the establishment of a central data center by some technical society such as the IEEE. Data processing firms could provide the center with information on errors, cost of errors, and the effectiveness of error reduction schemes. If necessary, the names of firms could be kept confidential in order to protect the trade secrecy of poor performance.

REFERENCES

1. J.H. Bauch, "Cut Input Costs With Key-To-Tape Devices," Computer Decisions, May 1971, pp. 36-39.
2. R.M. Davis, Letter to Hon. Richard T. Hanna of California, Congressional Record, November 19, 1971, pp. E12463.
3. G. Minton, Evaluation of Prequalification Training of Coders, and Some Factors Bearing on 2-Way vs 3-Way Inspection Schemes, Bureau of the Census, Statistical Research Division, April 1969.
4. A.H. Bieser, "Key-To-Tape and Key-To-Disc Systems," Product Profiles, Modern Data, December 1971.
5. L.A. Feidelman, "Keyboard-To-Tape Equipment Economics - Part 2," Source Data Automation, Modern Data, October 1970.
6. W.H. Von Alven, Reliability Engineering, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1964.
7. M. Klerer, G.A. Korn, Digital Computer User's Handbook, McGraw-Hill Book Company, 1967.
8. "Errors and Error Management," Communications Clinic, Modern Data, March 1972, pp. 22-24.
9. Alexander, Gryb, and Nast, Capabilities of the Telephone Network for Data Transmission," Bell System Technical Journal, Vol. 39, pg. 431, 1960.
10. Townsend & Watts, "Effectiveness of Error Control Data Communications Over the Switched Telephone Network," Bell System Technical Journal, Vol. 43, pg. 2611.
11. "1969-70 Connection Survey," A Compendium of Papers, Bell System Technical Journal, Vol. 50, April 1971.
12. J.J. Waterman, Jr., "Uninterruptible Power Systems," Modern Data, May 1972, pp. 30-35.
13. E. Yourdon, "Reliability of Real-Time Systems, Part 3, The Causes of System Failures," Modern Data, March 1972, pp. 36-41.
14. E. Yourdon, "Reliability of Real-Time Systems, Part 1, Different Concepts of Reliability," Modern Data, January 1972, pp. 36-42.
15. ASQC Standard A3-1964, Glossary of General Terms Used in Quality Control.
16. E.G. Kirkpatrick, Quality Control for Managers and Engineers, John Wiley & Sons, Inc., 1970.
17. D.H. Brandon, Management Standards for Data Processing, D. Van Nostrand Company, Inc., Princeton, N.J., 1963.
18. G. Minton, "Some Decision Rules for Administrative Applications of Quality Control," Journal of Quality Technology, Aprol 1970, pp. 86-98.
19. G. Minton, A.F. Krivitsky, "Two Additional Properties of a Model for Quality Control of Administrative Processes," Journal of Quality Technology, January 1971, pp. 6-17.
20. M. Marvin, "Inspection Efficiency and Sampling Inspection Plans," Journal of the American Statistical Association, Vol. 41, No. 236, December 1946, pp. 432-438.
21. U.S. Bureau of the Census. United States Censuses of Population and Housing 1960: Quality Control of Preparatory Operations, Microfilming, and Coding. Washington, D.C., 1965, pp. 57.
22. G. Minton, Formulas for Analyzing the Effect of Inspection and Correction Error on Sampling Inspection Plans in Data Processing Activities. U.S. Bureau of the Census, Statistical Research Division, January 1969.
23. D.K. Lloyd and M. Lipow, Reliability: Management, Methods, and Mathematics, Englewood Cliffs, New Jersey, Prentice Hall, Inc., 1962, pp. 287-291.
24. A. Wald, Sequential Analysis, New York, John Wiley & Sons, Inc., 1947, pp. 88-105.

TABLE 1

$Q_{d,s}$: PROBABILITY OF s SUCCESSIVE ACCEPTED SAMPLES IN MAXIMUM OF d INSPECTED[18]

L_p	$Q_{6,2}$	$Q_{6,3}$	$Q_{10,2}$	$Q_{10,3}$	$Q_{10,4}$	$Q_{10,5}$	$Q_{12,4}$	$Q_{12,5}$	$Q_{12,6}$
0.05	0.0096	0.0004	0.0214	0.0010	0.0000	0.0000	0.0001	0.0000	0.0000
0.10	0.0369	0.0028	0.0803	0.0073	0.0006	0.0001	0.0008	0.0001	0.0000
0.15	0.0794	0.0091	0.1675	0.0234	0.0031	0.0004	0.0039	0.0005	0.0001
0.20	0.1347	0.0208	0.2733	0.0523	0.0093	0.0016	0.0118	0.0021	0.0004
0.25	0.2002	0.0391	0.3882	0.0961	0.0215	0.0046	0.0272	0.0061	0.0013
0.30	0.2733	0.0648	0.5036	0.1551	0.0420	0.0109	0.0531	0.0143	0.0038
0.35	0.3516	0.0986	0.6126	0.2285	0.0731	0.0223	0.0919	0.0291	0.0090
0.40	0.4326	0.1408	0.7100	0.3141	0.1167	0.0410	0.1455	0.0531	0.0188
0.45	0.5141	0.1914	0.7927	0.4086	0.1740	0.0692	0.2147	0.0890	0.0357
0.50	0.5938	0.2500	0.8594	0.5078	0.2451	0.1094	0.2988	0.1394	0.0625
0.55	0.6697	0.3161	0.9102	0.6070	0.3293	0.1636	0.3957	0.2061	0.1024
0.60	0.7402	0.3888	0.9467	0.7014	0.4245	0.2333	0.5013	0.2897	0.1586
0.65	0.8036	0.4669	0.9710	0.7864	0.5272	0.3191	0.6103	0.3892	0.2338
0.70	0.8590	0.5488	0.9859	0.8586	0.6325	0.4202	0.7160	0.5015	0.3294
0.75	0.9053	0.6328	0.9941	0.9154	0.7347	0.5339	0.8116	0.6209	0.4449
0.80	0.9421	0.7168	0.9980	0.9562	0.8273	0.6554	0.8905	0.7392	0.5767
0.85	0.9693	0.7984	0.9995	0.9818	0.9039	0.7765	0.9481	0.8461	0.7166
0.90	0.9874	0.8748	0.9999	0.9948	0.9594	0.8857	0.9830	0.9306	0.8503
0.95	0.9971	0.9431	1.0000	0.9994	0.9909	0.9672	0.9977	0.9832	0.9556
1.00	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

L_p	$Q_{16,5}$	$Q_{16,6}$	$Q_{16,7}$	$Q_{20,5}$	$Q_{20,6}$	$Q_{20,8}$	$Q_{25,5}$	$Q_{25,8}$
0.05	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.10	0.0001	0.0000	0.0000	0.0001	0.0000	0.0000	0.0002	0.0000
0.15	0.0007	0.0001	0.0000	0.0010	0.0001	0.0000	0.0014	0.0000
0.20	0.0029	0.0005	0.0001	0.0042	0.0008	0.0000	0.0054	0.0000
0.25	0.0083	0.0019	0.0004	0.0119	0.0028	0.0002	0.0156	0.0002
0.30	0.0194	0.0053	0.0014	0.0278	0.0079	0.0006	0.0361	0.0008
0.35	0.0392	0.0126	0.0040	0.0558	0.0185	0.0020	0.0720	0.0027
0.40	0.0710	0.0262	0.0095	0.1001	0.0383	0.0054	0.1283	0.0073
0.45	0.1180	0.0492	0.0202	0.1642	0.0714	0.0128	0.2080	0.0174
0.50	0.1826	0.0854	0.0390	0.2499	0.1223	0.0273	0.3116	0.0369
0.55	0.2660	0.1383	0.0699	0.3559	0.1950	0.0534	0.4349	0.0716
0.60	0.3670	0.2110	0.1173	0.4777	0.2912	0.0967	0.5689	0.1284
0.65	0.4821	0.3049	0.1854	0.6064	0.4096	0.1635	0.7009	0.2138
0.70	0.6045	0.4191	0.2780	0.7308	0.5438	0.2594	0.8168	0.3319
0.75	0.7250	0.5487	0.3960	0.8385	0.6825	0.3867	0.9052	0.4805
0.80	0.8327	0.6845	0.5365	0.9196	0.8100	0.5412	0.9614	0.6470
0.85	0.9173	0.8127	0.6899	0.9700	0.9101	0.7084	0.9892	0.8070
0.90	0.9718	0.9165	0.8381	0.9933	0.9719	0.8618	0.9984	0.9288
0.95	0.9960	0.9808	0.9533	0.9996	0.9967	0.9668	1.0000	0.9897
1.00	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

TABLE 2

$S_{d,f}$: PROBABILITY OF f OR FEWER REJECTED SAMPLES WITHIN d INSPECTED[18]

L_p	$S_{5,1}$	$S_{5,2}$	$S_{10,1}$	$S_{10,2}$	$S_{10,3}$	$S_{10,4}$	$S_{15,1}$	$S_{15,2}$	$S_{15,3}$
0.05	0.0000	0.0012	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.10	0.0005	0.0086	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000
0.15	0.0022	0.0266	0.0000	0.0000	0.0001	0.0014	0.0000	0.0000	0.0000
0.20	0.0067	0.0579	0.0000	0.0001	0.0009	0.0064	0.0000	0.0000	0.0000
0.25	0.0156	0.1035	0.0000	0.0004	0.0035	0.0197	0.0000	0.0000	0.0000
0.30	0.0308	0.1631	0.0001	0.0016	0.0106	0.0473	0.0000	0.0000	0.0001
0.35	0.0540	0.2352	0.0005	0.0048	0.0260	0.0949	0.0000	0.0001	0.0005
0.40	0.0870	0.3174	0.0017	0.0123	0.0548	0.1662	0.0000	0.0003	0.0019
0.45	0.1312	0.4069	0.0045	0.0274	0.1020	0.2616	0.0001	0.0011	0.0063
0.50	0.1875	0.5000	0.0107	0.0547	0.1719	0.3770	0.0005	0.0037	0.0176
0.55	0.2562	0.5931	0.0233	0.0996	0.2660	0.5044	0.0017	0.0107	0.0424
0.60	0.3370	0.6826	0.0464	0.1673	0.3823	0.6331	0.0052	0.0271	0.0905
0.65	0.4284	0.7648	0.0860	0.2616	0.5138	0.7515	0.0142	0.0617	0.1727
0.70	0.5282	0.8369	0.1493	0.3828	0.6496	0.8497	0.0353	0.1268	0.2969
0.75	0.6328	0.8965	0.2440	0.5256	0.7759	0.9219	0.0802	0.2361	0.4613
0.80	0.7373	0.9421	0.3758	0.6778	0.8791	0.9672	0.1671	0.3980	0.6482
0.85	0.8352	0.9734	0.5443	0.8202	0.9500	0.9901	0.3186	0.6042	0.8227
0.90	0.9185	0.9914	0.7361	0.9298	0.9872	0.9984	0.5490	0.8159	0.9444
0.95	0.9774	0.9988	0.9139	0.9885	0.9990	0.9999	0.8290	0.9638	0.9945
1.00	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

L_p	$S_{15,4}$	$S_{20,1}$	$S_{20,2}$	$S_{20,3}$	$S_{21,4}$	$S_{20,4}$	$S_{20,5}$	$S_{20,6}$
0.05	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.20	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.25	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.30	0.0007	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.35	0.0028	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.40	0.0093	0.0000	0.0000	0.0000	0.0003	0.0000	0.0000	0.0001
0.45	0.0255	0.0000	0.0000	0.0003	0.0015	0.0000	0.0000	0.0001
0.50	0.0592	0.0000	0.0002	0.0013	0.0059	0.0000	0.0002	0.0007
0.55	0.1204	0.0001	0.0009	0.0049	0.0189	0.0002	0.0011	0.0040
0.60	0.2173	0.0005	0.0036	0.0160	0.0510	0.0015	0.0057	0.0172
0.65	0.3519	0.0021	0.0121	0.0444	0.1182	0.0075	0.0233	0.0586
0.70	0.5155	0.0076	0.0355	0.1071	0.2375	0.0302	0.0766	0.1595
0.75	0.6865	0.0243	0.0913	0.2252	0.4148	0.0979	0.2026	0.3481
0.80	0.8358	0.0692	0.2061	0.4114	0.6296	0.2552	0.4275	0.6070
0.85	0.9383	0.1756	0.4049	0.6477	0.8298	0.5245	0.7106	0.8474
0.90	0.9873	0.3917	0.6769	0.8670	0.9568	0.8245	0.9268	0.9742
0.95	0.9994	0.7358	0.9245	0.9841	0.9974	0.9844	0.9967	0.9994
1.00	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

TABLE 3

$U_{D,C}$: PROBABILITY OF SURVIVING D DECISIONS GIVEN AN INITIAL BONUS OF C POINTS[19]

L_p	$U_{5,1}$	$U_{5,2}$	$U_{10,1}$	$U_{10,2}$	$U_{10,3}$	$U_{10,4}$	$U_{10,5}$	$U_{15,1}$	$U_{15,2}$	$U_{15,3}$	$U_{15,4}$	$U_{15,5}$	$U_{20,3}$	$U_{20,4}$
0.05	0.0006	0.0118	0.0000	0.0000	0.0005	0.0008	0.0075	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.10	0.0044	0.0442	0.0003	0.0009	0.0061	0.0103	0.0481	0.0000	0.0001	0.0002	0.0010	0.0017	0.0000	0.0000
0.15	0.0140	0.0933	0.0020	0.0056	0.0255	0.0410	0.1292	0.0002	0.0010	0.0021	0.0082	0.0130	0.0004	0.0008
0.20	0.0310	0.1552	0.0073	0.0190	0.0656	0.1009	0.2424	0.0011	0.0055	0.0108	0.0321	0.0483	0.0035	0.0061
0.25	0.0566	0.2266	0.0189	0.0466	0.1295	0.1904	0.3734	0.0047	0.0189	0.0352	0.0840	0.1203	0.0157	0.0260
0.30	0.0913	0.3042	0.0398	0.0926	0.2158	0.3032	0.5072	0.0144	0.0482	0.0847	0.1701	0.2319	0.0481	0.0754
0.35	0.1348	0.3853	0.0722	0.1587	0.3195	0.4288	0.6319	0.0349	0.0996	0.1655	0.2877	0.3732	0.1122	0.1661
0.40	0.1869	0.4672	0.1175	0.2437	0.4329	0.5555	0.7393	0.0705	0.1763	0.2760	0.4255	0.5258	0.2136	0.2979
0.45	0.2465	0.5478	0.1759	0.3432	0.5477	0.6728	0.8257	0.1243	0.2763	0.4074	0.5676	0.6698	0.3469	0.4562
0.50	0.3125	0.6250	0.2461	0.4512	0.6563	0.7734	0.8906	0.1964	0.3928	0.5455	0.6982	0.7899	0.4966	0.6167
0.55	0.3835	0.6973	0.3257	0.5603	0.7523	0.8534	0.9361	0.2836	0.5156	0.6754	0.8062	0.8789	0.6423	0.7563
0.60	0.4579	0.7632	0.4117	0.6639	0.8320	0.9122	0.9657	0.3804	0.6339	0.7855	0.8865	0.9376	0.7670	0.8613
0.65	0.5341	0.8218	0.5004	0.7561	0.8938	0.9520	0.9833	0.4803	0.7389	0.8697	0.9401	0.9716	0.8614	0.9299
0.70	0.6105	0.8722	0.5885	0.8333	0.9383	0.9765	0.9929	0.5776	0.8252	0.9280	0.9720	0.9889	0.9251	0.9688
0.75	0.6855	0.9141	0.6730	0.8941	0.9678	0.9900	0.9974	0.6682	0.8910	0.9643	0.9887	0.9964	0.9635	0.9880
0.80	0.7578	0.9472	0.7518	0.9387	0.9854	0.9965	0.9993	0.7503	0.9378	0.9845	0.9962	0.9991	0.9844	0.9961
0.85	0.8260	0.9718	0.8239	0.9690	0.9946	0.9991	0.9999	0.8236	0.9689	0.9945	0.9990	0.9998	0.9944	0.9989
0.90	0.8894	0.9882	0.8889	0.9877	0.9986	0.9998	1.0000	0.8889	0.9877	0.9986	0.9998	1.0000	0.9986	0.9998
0.95	0.9474	0.9973	0.9474	0.9972	0.9999	1.0000	1.0000	0.9474	0.9972	0.9999	1.0000	1.0000	0.9996	0.9999
1.00	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

TABLE 4

$C_{D,C}$: EXPECTED NUMBER OF CREDITS, GIVEN CONTINUANCE IN PROCESS AT END OF D DECISIONS[19]

L_p	$1 - L_p$	$C_{5,1}$	$C_{5,2}$	$C_{10,1}$	$C_{10,2}$	$C_{10,3}$	$C_{10,4}$	$C_{10,5}$	$C_{15,1}$	$C_{15,2}$	$C_{15,3}$	$C_{15,4}$	$C_{15,5}$	$C_{20,3}$	$C_{20,4}$
0.05	0.95	2.083	1.183	1.220	2.132	1.238	2.155	1.275	2.149	1.249	2.158	1.264	2.176	1.264	2.174
0.10	0.90	2.172	1.372	1.453	2.280	1.492	2.334	1.583	2.319	1.519	2.340	1.554	2.384	1.561	2.378
0.15	0.85	2.268	1.568	1.701	2.446	1.766	2.539	1.920	2.513	1.813	2.550	1.874	2.628	1.887	2.618
0.20	0.80	2.371	1.771	1.965	2.633	2.063	2.773	2.292	2.736	2.136	2.795	2.232	2.916	2.254	2.903
0.25	0.75	2.483	1.983	2.248	2.844	2.387	3.043	2.706	2.994	2.494	3.078	2.634	3.260	2.670	3.246
0.30	0.70	2.604	2.204	2.554	3.084	2.741	3.355	3.169	3.293	2.893	3.413	3.092	3.672	3.149	3.661
0.35	0.65	2.734	2.434	2.884	3.358	3.131	3.717	3.686	3.642	3.342	3.809	3.616	4.168	3.706	4.168
0.40	0.60	2.877	2.677	3.243	3.669	3.562	4.135	4.263	4.051	3.851	4.279	4.220	4.764	4.358	4.789
0.45	0.55	3.031	2.931	3.635	4.026	4.040	4.618	4.905	4.530	4.430	4.837	4.920	5.480	5.129	5.552
0.50	0.50	3.200	3.200	4.064	4.433	4.571	5.172	5.614	5.092	5.092	5.499	5.729	6.330	6.042	6.486
0.55	0.45	3.384	3.484	4.534	4.898	5.162	5.803	6.390	5.748	5.848	6.281	6.660	7.324	7.120	7.618
0.60	0.40	3.585	3.785	5.050	5.427	5.817	6.513	7.228	6.510	6.710	7.193	7.719	8.462	8.380	8.961
0.65	0.35	3.805	4.105	5.617	6.025	6.538	7.301	8.120	7.385	7.685	8.241	8.904	9.728	9.829	10.510
0.70	0.30	4.045	4.445	6.238	6.696	7.327	8.157	9.056	8.374	8.774	9.420	10.200	11.096	11.453	12.237
0.75	0.25	4.308	4.808	6.913	7.440	8.178	9.071	10.022	9.471	9.971	10.712	11.585	12.534	13.221	14.094
0.80	0.20	4.595	5.195	7.643	8.254	9.084	10.026	11.007	10.662	11.262	12.093	13.030	14.009	15.094	16.031
0.85	0.15	4.907	5.607	8.424	9.126	10.032	11.007	12.001	11.928	12.628	13.533	14.508	15.502	17.033	18.008
0.90	0.10	5.246	6.046	9.250	10.050	11.008	12.001	13.000	13.250	14.050	15.008	16.001	17.000	19.008	20.001
0.95	0.05	5.611	6.511	10.111	11.011	12.001	13.000	14.000	14.611	15.511	16.501	17.500	18.500	21.001	22.000
1.00	0.00	6.000	7.000	11.000	12.000	13.000	14.000	15.000	16.000	17.000	18.000	19.000	20.000	23.000	24.000

TABLE 5

$Z_{D,C}$: EXPECTED NUMBER DECISIONS, GIVEN REMOVAL ON OR BEFORE D DECISIONS[19]

L_p	$1-L_p$	$Z_{5,1}$	$Z_{5,2}$	$Z_{10,1}$	$Z_{10,2}$	$Z_{10,3}$	$Z_{10,4}$	$Z_{10,5}$	$Z_{15,1}$	$Z_{15,2}$	$Z_{15,3}$	$Z_{15,4}$	$Z_{15,5}$	$Z_{20,3}$	$Z_{20,4}$
0.00	1.00	1.000	2.000	1.000	2.000	3.000	4.000	5.000	1.000	2.000	3.000	4.000	5.000	3.000	4.000
0.05	0.95	1.108	2.174	1.111	2.222	3.330	4.438	5.511	1.111	2.222	3.333	4.444	5.555	3.333	4.444
0.10	0.90	1.221	2.305	1.247	2.491	3.700	4.917	5.960	1.250	2.499	3.748	4.988	6.230	3.750	4.999
0.15	0.85	1.332	2.406	1.406	2.798	4.070	5.376	6.312	1.426	2.842	4.256	5.615	6.990	4.278	5.700
0.20	0.80	1.433	2.485	1.581	3.120	4.408	5.774	6.578	1.647	3.249	4.840	6.267	7.748	4.935	6.555
0.25	0.75	1.522	2.546	1.759	3.429	4.695	6.098	6.775	1.910	3.692	5.448	6.875	8.424	5.692	7.506
0.30	0.70	1.595	2.592	1.925	3.703	4.926	6.349	6.920	2.196	4.128	6.016	7.391	8.974	6.472	8.438
0.35	0.65	1.653	2.625	2.066	3.926	5.101	6.534	7.022	2.473	4.512	6.492	7.792	9.388	7.175	9.239
0.40	0.60	1.694	2.649	2.174	4.089	5.223	6.661	7.091	2.704	4.807	6.844	8.074	9.673	7.722	9.839
0.45	0.55	1.719	2.662	2.242	4.188	5.294	6.734	7.130	2.857	4.992	7.059	8.241	9.839	8.065	10.204
0.50	0.50	1.727	2.667	2.264	4.221	5.318	6.759	7.143	2.910	5.055	7.131	8.296	9.893	8.181	10.327
0.55	0.45	1.719	2.662	2.242	4.188	5.294	6.734	7.130	2.857	4.992	7.059	8.241	9.839	8.065	10.204
0.60	0.40	1.694	2.649	2.174	4.089	5.223	6.661	7.091	2.704	4.807	6.844	8.074	9.673	7.722	9.839
0.65	0.35	1.653	2.625	2.066	3.926	5.101	6.534	7.022	2.473	4.512	6.492	7.792	9.388	7.175	9.239
0.70	0.30	1.595	2.592	1.925	3.703	4.926	6.349	6.920	2.196	4.128	6.016	7.391	8.974	6.472	8.438
0.75	0.25	1.522	2.546	1.759	3.429	4.695	6.098	6.775	1.910	3.692	5.448	6.875	8.424	5.692	7.506
0.80	0.20	1.433	2.485	1.581	3.120	4.408	5.774	6.578	1.647	3.249	4.840	6.267	7.748	4.935	6.555
0.85	0.15	1.332	2.406	1.406	2.798	4.070	5.376	6.312	1.426	2.842	4.256	5.615	6.990	4.278	5.700
0.90	0.10	1.221	2.035	1.247	2.491	3.700	4.917	5.960	1.250	2.499	3.748	4.988	6.203	3.749	4.999
0.95	0.05	1.108	2.174	1.111	2.222	3.330	4.438	5.511	1.111	2.222	3.333	4.444	5.555	3.333	4.444

TESTING OF SPACECRAFT IN LONG-TERM STORAGE

R. J. Masterson and R. N. Miller
TRW Systems Group
Redondo Beach, California

SUMMARY

This paper addresses the problem of developing a storage and test policy which may be applied to equipment placed in long-term storage prior to utilization. A closed-form analytic solution is developed to aid in evaluating the characteristics of a given policy in terms of test efficiency, on time delivery, and subsequent reliable operation of the equipment in its intended application. The analysis model is, to the authors knowledge, new and hence is described in detail; a computer program based on the model is outlined and a flow diagram of its logic included. Problems associated with estimating valid input data are treated. An example of the application of the analysis to spacecraft is presented to fully illustrate the approach. Finally, the paper closes with a discussion of some of the many situations in which such a tool could be employed in a variety of industrial, research, and sports contexts.

1. INTRODUCTION

The advent of replenishable multi-satellite systems in recent years (TIROS is a good example) has created a requirement for the long-term storage of spacecraft. Such a requirement led to the development of the analysis tool described in this paper. The problems will vary from case to case depending largely on two types of factors: (i) the engineering characteristics of the hardware involved (e.g., susceptibility to corrosion, sensitivity to a 1.0g field, potential temperature effects, etc.); and (ii) the characteristics of the mission which the hardware is called upon to perform (e.g., is the system repairable or not after it is placed in operation?, is the demand for the equipment random or based upon a predetermined schedule?, how soon after usage demand must the hardware be put into service?, etc.). This paper deals largely with the second class of decisions which must be approached in many cases on a statistical basis. The first class of problems is more deterministic in nature, and generally fairly well understood in the industry: storage in dry nitrogen with periodic rotation of 1.0g sensitive components are among widely followed policies.

While the illustrations included in the paper are written largely in the context of a stored spacecraft subject to an unscheduled launch call, the same problem occurs for other classes of equipment; notably weapon systems held in readiness for use only in the event of an emergency, electronic parts stored prior to assembly, TV sets stored in a warehouse or showroom, and used cars on the corner lot. The primary concern with stored equipment is that it works properly when it is called upon to be used. This motivation usually results in some testing being performed on the equipment after it is taken out of storage prior to its actual use. Problems uncovered during the post call-up tests are repaired prior to use to provide maximum confidence that a successful mission will result. Another factor important in many operational contexts is a requirement to respond very rapidly to the activation call-up; to compensate for reduced testing after call-up, periodic testing during the storage period is often considered. This paper describes a dynamic model of this situation which may be exercised to define a storage and test policy consistent with the dual objectives of (i) responding on time to the call-up, and (ii) having adequate assurance that the equipment is failure free at the time its utilization begins.

2. ANALYSIS MODEL

The Storage Sequence of Events

The general sequence of events occurring in a typical storage situation is illustrated in Figure 1. The equipment is placed into storage; at that time there may be undetected failures present. These failures could result from inadequate testing or the inability of the test equipment to detect the failure. Some storage time occurs and depending upon the individual storage policy being evaluated, testing may be done on a periodic basis. Failures may occur during storage, during test or may be induced by the testing itself. The test will detect some percentage (typically not all) of whatever defects are present (including previously undetected defects) before the equipment is returned to storage. Finally, the call to use the stored item occurs, a final test may or may not be done and the equipment is then applied to its end use, hopefully free of failures (no undetected failures).

The model described herein is a probabilistic analysis of the various events which take place in the storage sequence. It uses an analytic of closed-form approach as contrasted with a Monte Carlo simulation approach; consequently, the computer run time required to evaluate an individual case is very nominal. The direct thrust of the model output focuses on (i) the number of failures which may be detected after the call-up decision has been made and (ii) the consequent delay induced by these detected failures assuming that usage will not begin until all anomalies are repaired and retest completed. As was indicated in the summary, however, almost any type of information describing the effectiveness of a storage policy may be determined. Through a suitable structuring of multiple cases, one may estimate (i) the efficiency of a test (how many failures it detects versus how many it introduces); (ii) the number of defects present when storage begins; and (iii) the number of defects likely to remain when the equipment is put into service. These points are further illustrated in the examples discussed in Section 3 of this paper. The computer program which implements this analysis thus represents a tool which may be used in a variety of ways to evaluate many aspects of the storage problem. As with any tool, the way in which it will be applied depends upon the job to be done.

Computer Program Inputs and Outputs

The computer program requires inputs which describe the pertinent characteristics of the hardware and of the storage policy being contemplated for that hardware. In addition, certain input variables are defined solely in the interests of computer processing efficiency and output format standardization. The direct outputs of the program describe the number of failures detected after the use decision, and the days of delay induced by these failures before the equipment is available for actual service. As discussed

earlier, many other conclusions may be drawn based upon the joint evaluation of several structured cases.

The basic inputs and outputs are listed in Table 1. The launch call time is treated as an input variable; by varying launch call time, a given test schedule can be evaluated against a range of contingencies. Many of the inputs are accumulated into a data file made up of data which remains largely invariant for a large number of cases. These include the characteristics of the hardware itself for each item: its likelihood of failure in various tests; its storage failure rate; the likelihood failures will be detected for each item in various tests; the probability each item will have undetected failures when entering storage; and the repair time in days should an item fail. Other inputs which vary from case to case are input directly for each case.

Table 1. Analysis Model Inputs and Outputs

INPUTS

- Call-up time and call-up test type.
- Number and types of tests scheduled prior to call-up.
- Interval between tests.
- Storage failure rates for each item (or subcategory of equipment).
- Probability of no defects present when put into storage for each item.
- Restoration times should failure occur during testing for each item.
- For each type of test:
  - Probability of no additional failures during test for each item,
  - Probability failure present before test will be detected during test for each item,
  - Probability new failure occurring during test will be detected during test for each item.
- Complexity factor (a measure of number of failures anticipated, used only to size matrices in computer program).
- Units factor (delay measured in days of some multiple thereof).
- Efficiency factor (measure of non-additivity of actual delays, amount of multiple repairs being simultaneously performed).

OUTPUTS

- Probability distribution of number of detected failures for each item during call-up test sequence.
- Probability distribution of days of delay due to failures detected during call-up test sequence.
- Average days of delay due to failures detected during launch call-up test sequence.

Steps in Analysis

The analysis of the storage process shown in Figure 1 is implemented through the steps described in Figure 2. These steps also correlate with sections of the computer program. The method of analysis operates upon probability distributions which govern the various events. Some overall assumptions about the form of these distributions built into the analysis include:

- Storage failure rates are constant with time; hence, the number of failures, F, occurring in storage follows a Poisson distribution,

$$P(F=x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}, \quad x = 0, 1, 2, \ldots$$

where,

λ = in-storage failure rate
t = time in storage
F = number of failures

- Number of failures occurring during test also is treated with Poisson distribution; also number of defects present prior to storage.
- An item is spared or unspared; for items which are spared it is assumed that a spare is available in the event of test failure.

Figure 2 is a simplified flow diagram of these steps. A careful review of Table 2 and Figure 2 should result in the development of an adequate understanding of the analysis so that the reader could apply the methology with his own computer program. The method is further illustrated by the examples contained in Section 3. Another approach to modeling a similar situation using Markov chains may be found in Reference 1.

Table 2. Steps in Analysis and Computer Program

1. Input complexity factor, units factor, efficiency factor, E'.
2. Input time to call-up in days, type of call-up test.
3. Input number of scheduled tests prior to call-up.
4. Dimension matrices.
5. Input test type for each scheduled test.
6. Input storage time between each scheduled test.
7. Read data file.
8. N = Number of subsystems or items, N1 = Number of tests prior to use.
9. Start with first item.
10. W = Probability of at least one defect entering storage.
11. Determine distribution of number of defects present upon entering storage:

    $W(0) = 1 - W$

    Thus, letting $v = -\log(1-W)$, the Poisson parameter consistent with $W(0) = 1 - W$, one obtains

    $$W(y) = \frac{e^{-v} v^y}{y!}, \quad y = 0, 1, 2, \ldots$$

12. Determine distribution of number of defects occurring in storage prior to first test.

    $$R(x) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}, \quad x = 0, 1, 2, \ldots$$

    where λ = in-storage failure rate
    t = time to start of first test

13. Convolute R and W to obtain distribution of total defects present upon entering first test.

    $$B(x) = \sum_{k=0}^{x} R(k)\, W(x-k), \quad x = 0, 1, 2, \ldots$$

    B is the distribution of R+W.

14. D = Probability defect entering test will be detected.

15. Determine matrix D(M1, M2) the probability that M2 of M1 defects will be detected.

$$D(M1, M2) = \binom{M1}{M2} D^{M2}(1-D)^{M1-M2}, \; M1=0, 1, 2, \ldots, \quad M2=0, 1, 2, \ldots, M1$$

using the standard binomial distribution for M2 successes out of M1 trials.

16. X = Probability of at least one defect occurring during test.

17. Determine distribution of number of defects occurring during test:

$X(0) = 1-X$

Letting a = -log (1-X), one obtains

$$X(x) = \frac{e^{-a} a^x}{x!}, \; x = 0, 1, 2, \ldots$$

18. E = Probability of detecting failure occurring during test.

19. Determine matrix E(M1, M2):

$$E(M1M2) = \binom{M1}{M2} E^{M2} (1-E)^{M1-M2}, \; M1=1,2,\ldots,; \quad M2=0,1,2,\ldots,M1 \text{ similar to D.}$$

20. Determine distribution of number of failures present prior to test which go undetected:

$$N(x) = \sum_{k=x}^{\infty} B(k)\, D(k,k-x), \; x = 0, 1, 2, \ldots$$

21. Determine distribution of number of failures occurring during test which go undetected:

$$N\$(x) = \sum_{k=x}^{\infty} X(k)\, E(k,k-x), \; x = 0,1,2,\ldots$$

22. Convolute N and N$ to obtain distribution of total defects which are present but undetected at end of test:

$$W(x) = \sum_{k=0}^{x} N(k)\, N\$(x-k), \; x = 0, 1, 2, \ldots$$

23. Proceed to next test using W from 22 as the distribution of defects present prior to next test. Repeat steps 12 through 22 for next test.

24. Repeat 23 until last test prior to call-up is completed, test Number N1-1.

25. Repeat steps 12 through 19 for call-up test.

26. Determine distribution of number of failures present prior to call-up test which are detected:

$$N(x) = \sum_{k=x}^{\infty} B(k)\, D(k,x), \; x = 0, 1, 2, \ldots$$

27. Determine distribution of number of failures occurring during call-up test which are detected:

$$N\$(x) = \sum_{k=x}^{\infty} X(k)\, E(k,x), \; x = 0, 1, 2, \ldots$$

28. Convolute N and N$ to obtain distribution of total defects which are detected during call-up test:

$$W(x) = \sum_{k=0}^{w} N(k)\, N\$(x-k), \; x = 0, 1, 2, \ldots$$

29. Print out distribution of number of detected failures for item number one. Also determine and print out average number of detected failures.

30. Determine Z(y), the probability of y days delay due to item number one:

$$Z(y) = \sum_{k \varepsilon M} W(k), \; y = 0, 1, 2, \ldots \text{ in days}$$

where $M = (k: y < k \cdot T' \cdot E' \leq y+1)$

T' = repair time for item for one detected defect.

Probability associated with repair times greater than y days but less than y+1 days are grouped as the probability of a y day delay due to item number one.

31. Repeat steps 10 through 29 for item number two.

32. Determine Z$(y), the probability of y days delay due to item two:

$$Z\$(y) = \sum_{k \varepsilon M} w(k), \; y = 0, 1, 2, \ldots \text{ in days}$$

where $M = (K: y < k \cdot T' \cdot E' \leq y+1)$

33. Convolute Z and Z$ to obtain distribution of the probability of y days delay due to items one and two:

$$Y(y) = \sum_{k=0}^{y} Z(k) \cdot Z\$(y-k), \; y = 0, 1, 2, \ldots$$

34. Let Z(y) = Y(y) for y = 0, 1, 2, . . .

35. Repeat steps 10 through 29 and steps 32 through 34 for items three, four, ..., N.

36. Print out average days delay for each item and percent of total each contributes.

37. Print out final Z(y), y = 0, 1, 2, ..., the distribution of days delay due to items one through N.

3. EXAMPLE OF SPACECRAFT APPLICATION

This section illustrates the application of the analysis to the storage of a spacecraft while awaiting an unscheduled launch call. The manner in which input data were estimated, the way cases were structured, and the types of conclusions which were drawn are all described.

Data Requirements

The inputs to the model can be determined from a variety of sources. If the equipment under consideration is mature (e.g., a TV set), then the probability of failure could be determined using the failure rate and the operating time of the test. Another method would be the failure history of the equipment during tests similar to those to be conducted during storage and reactivation. If the equipment is stored in the unpowered state, References 2, 3, and 4 provide data for determining the probability of failure during storage.

The probability of detecting falures (test efficiency) is a function of the parameters tested, the environment of the test (ambient, hot, cold) and the failure rates of the test equipment (probability of not detecting a failure when it occurs). A method of determining test eeficiency is to compare, for each test considered, those parameters tested versus those not tested. This results in a percentage of the total parameters tested and allows for a comparison of the efficiencies of the various tests; however, the environments and test equipment must still be accounted for. It is also possible to determine the test efficiency from the failure data.

For the example shown herein, it was decided to use the failure data collected during spacecraft testing. It was felt that this single data source accounted for all aspects of determining the probability of

detecting failures. The two inputs that differed from this were the storage failure rate and the repair time. The storage failure rate used was taken to be 10% of the predicted operating failure rate. The repair time was estimated based on experience from troubleshooting, removing, replacing, and retest for the spacecraft under consideration.

Each failure report was reviewed and classified as to type of failure, type of test, subsystem and environment where the failure could be detected. The types of failures were as follows:

- Spacecraft Hardware Failure - This category was further subdivided to identify latent failures. These were workmanship failures attributed to manufacturing or a vendor that should (or could) have been detected during earlier testing.
- Test Failure - This category included all test equipment failures, procedure problems, operator error, etc.
- Non-Failure - These were usually minor out-of-tolerance conditions which did not result in replacement of hardware and were dispositioned "use-as-is." These were deleted from the sample since they did not require repair.

The types of testing were identified because the same types of testing as accomplished during integration and test were proposed for storage and reactivation. This allowed for the determination of failure probability and detection for the tests involved. Three general areas of testing were identified: subsystem integration, ambient test, and thermal vacuum. The ambient testing was further subdivided into integration system testing (IST), pre-thermal vacuum, and post-thermal vacuum.

The subsystems were those normally identified with a spacecraft; however, in some instances it was necessary to regroup some of the hardware into different categories because of differences in repair time since the program will only accept one repair time per subsystem. All the test failures were grouped into a subsystem identified as Test Equipment. This category also accounted for the possibility that test equipment could erroneously indicate the presence of a failure when none was present. The "subsystems" correspond to the "items" described earlier.

The identification of the various environments where the failure could be detected was necessary because some failures could only be detected at thermal vacuum conditions while others could only be detected at ambient conditions (visual inspection, etc.). For the most part, the failures could be detected at either environment.

Upon completion of the data review it was decided to eliminate those faillures occurring during subsystem integration since the objective was to determine the failures expected of a completely assembled spacecraft. This left a total of 12 ambient/thermal vacuum cycles upon which to base the probabilities required by the analysis model.

The task of determining failure probabilities from number of failures was accomplished by using the average number of failures detected and the Poisson distribution. The method is illustrated for a typical subsystem using the 12 ambient/thermal vacuum tests noted above.

a. Total number of failures detected - 28

b. Average failures per test $(\frac{28}{12})$ - 2.33

c. Probability of failure occurring:

$$P = 1-e^{-x} = 1-e^{-2.33} = 1-0.097$$

$$P = 0.903$$

Each subsystem was handled in the same manner. If no failure occurred in the subsystem, one failure was conservatively assumed. The probability of a defect existing when entering storage was determined in the same manner as above except the latent failures identified earlier were used.

The probability of detecting failures was determined from the latent failures and the environment where detected. As an example, one subsystem showed 10 latent failures which could be detected at either ambient or thermal vacuum. Of these, eight were detected at ambient and two at thermal vacuum; therefore, the probability of detecting latent failures at ambient was 80% and the probability for thermal vacuum was 20% better or 96%. The probability of detecting a failure occurring during test was assumed to be the same.

All subsystems did not display the above trend since latent failures were not prevalent and/or total failures were small and detected at both environments. It was assumed the detection capabilities would be the same regardless of the type of testing and would be at least as efficient as the most effective test (.96).

Results

A study was initiated with the purpose of evaluating an existing test plan for the long-term storage of a spacecraft which must be launched within 75 to 140 days after an unscheduled call-up. The difference in launch schedule resulted from the type of reactivation testing conducted prior to shipment (ambient versus ambient plus environmental testing). The original purpose of the study was to determine whether or not the recommended shipment dates could be met (30 days prior to launch). The analysis model developed for this study was directed at the number of days delay due to failures (and their repair times) in the system and had to be compared to the contingency allowed for these failures.

The test plan consisted of three different types of reactivation testing based on the time since the last thermal vacuum test (T/V). The least of the tests was an integrated system test (IST), the second was a very detailed ambient test and the final test was the ambient test plus a very difficult thermal vacuum (T/V) test. In addition the test plan called for the spacecraft to be stored without power applied, with the Propulsion Subsystem pressurized, a controlled humidity and a nitrogen environment.

In-storage testing was to consist of a quarterly electrical test which was basically an "aliveness test." Additionally a detailed ambient test, a T/V test and a post T/V ambient test were to be conducted at nine month intervals.

The analysis model was exercised for a variety of conditions to arrive at conclusions and recommendations for the test program under consideration. Table 3 is a sample of the cases conducted upon which the final conclusions and recommendations resulting from the study were based.

The first three cases were a variation in the launch call to exercise each of the three reactivation test schedules contained in the test plan. Case Number 1 shows the minimum reactivation testing; based on the failures and days delay it was considered to be a useless test and was eliminated from further consideration.

Knowing that the in-storage test was basically an "aliveness test," it was decided to study elimination of these tests to determine their effect on the overall program. Case 4 of Table 3 shows that they have no effect on either the days delay or the number of failures detected. As a result, these tests were also eliminated from further cosideration.

Table 3. Sample of Storage Study Considerations

Case Number	Launch Call (Mo.)	Tests During Storage (1)	Reactivation Testing	Average Failures Detected	Average Days Delay
1	12	1,1,2	IST	3.2	9.3
2	18	1,1,2, 1,1	T/V	13.1	28.8
3	24	1,1,2, 1,1,2	Ambient	7.7	18.7
4	24	2,2	Ambient	7.7	18.7
5	24	None	Ambient	9.9	22.8
6	24	None	T/V	14.1	30.7
7	24	None	T/V (2)	4.2	8.6
8	24	None	Ambient (3)	13.1	30.6
9	24	None	T/V (3)	15.9	38.4
10	24	2	T/V	11.9	26.4
11	24	2,2	T/V	11.3	25.2
12	24	2,2,2	T/V	11.3	25.1
13	24	2	T/V (3)	13.4	29.9
14	24	2,2	T/V (3)	12.7	28.5
15	24	2,2,2	T/V (3)	12.7	28.5
16	24	None	T/V	14.1	80.3

(1) Test Number 1 is an in-storage electrical (aliveness) test. Test Number 2 is an ambient plus a T/V test.

(2) Indicates that no failures occur during reactivation testing.

(3) Indicates perfect detection during reactivation testing.

The next step was to evaluate the effect of no testing during storage. Cases 5 and 6 study these effects. Both the failures and the delay were increased over Cases 2 and 3; however, these increases were not considered significant. The test plan contained 17 days contingency for the ambient test and 22 days for the T/V test. Since the test plan was based on a 40-hour single shift work week, the delays estimated by the analysis model were not considered excessive.

At this point a question was raised concerning the source of the failures; e.g., how many were being introduced by the testing and how many were present at the beginning of the test. Additionally, concern was shown over the efficiency of the tests and how many failures would still be remaining in the system after the reactivation testing, i.e., at launch. Cases Number 7, 8 and 9 are examples of runs conducted to answer the above questions. By changing the working copy of the program at the terminal, Case Number 7 was accomplished allowing no failures to occur during reactivation testing; comparing Cases 6 and 7 indicates that at least 4.2 failures were present at the start of test and 9.9 were caused by the test (14.1-4.2=9.9). The failures introduced by the tests were further identified as test or spacecraft failures. Some spacecraft failure were also identified as incipient failures accelerated by the T/V tests. Cases Number 8 and 9 were created by changing the program to allow perfect detection during reactivation testing. Comparing them with Cases 5 and 6 shows that failures will be undetected and remain in the system at launch. Also by dividing the failures in Cases 5 and 6 by those in Cases 8 and 9 it is seen that the T/V test is the most efficient even though more failures are introduced. (ambient is 76% effective while T/V is 89% effective). As a result, only T/V testing was considered in the remainder of the cases.

Cases 10 through 12 of the table represent an attempt to determine whether or not more intensive testing during the storage period would eliminate the failures remaining in the spacecraft. These three cases when compared to Case 6, indicated that some of the failures could be eliminated by interim testing; however, Cases 13 through 15 indicate that some failures are left in the system and cannot be eliminated (e.g., Cases 12 and 15 - 12.7-11.3=1.4). Since this matched our launch experience, the study was terminated at this point. It should also be pointed out that the failures used in the study were reviewed to determine their effect on orbit. For the most part, these failures were considered minor; this also coincided with past experience.

Up to this point in time, the spares complement was considered unlimited; e.g., it was assumed that an on-going program was following or that several vehicles were in storage allowing for units to be removed from other spacecraft. During the study, units began failing which had not failed previously and for which no spares were available. This resulted in a reevaluation of the spares complement and their recycle time through manufacturing, test, and return to the spacecraft. Case Number 16 is a sample of the days delay associated with this type of situation and represents an unacceptable condition.

Based upon the study results represented by Table 3, the conclusion was to place the spacecraft in storage, conduct no tests until the launch call is received and then to conduct a T/V test. Additionally, it was recommended that a complete complement of spares be provided. Management took the above results and, considering cost, manpower requirements, etc., arrived at the conclusion that interim T/V tests should be conducted at one year intervals for crew training purposes. They also recommended that the spares complement be increased over the present level.

Our customer has not, to date, directed us to revise our test program per the recommendations. He has, however, issued a contract to complete our spares complement.

4. OTHER APPLICATIONS

Other potential applications of the analysis tool described in this paper have been alluded to, things such as stored weapon system held in readiness in case of an emergency; piece parts placed in electronic stores prior to assembly; electronic boxes stored prior to integration into a system; TV sets and hi-fi equipment stored in warehouses before shipment to showrooms; used cars sitting on a lot waiting for a new owner to arrive. The list of such applications is potentially a very long one. This section examines a few such situations in more detail, pointing out some of the important questions which could be studied by use of such analyses. Two less obvious examples relating to psychology and baseball are touched upon.

Missiles in Silos

The equipment constituting the U.S. retaliatory strike capability is a classic example of stored equipment subject to a random or non-predetermined call-up use. The problems of what storage conditions, what

type of checkout procedures to apply and how often, have received extensive study over a long period of time. Analyses of the type discussed in the preceding sections should have played a key role in the reaching of decisions on these matters. Indeed, the need for nearly instantaneous response to the call-up has led to very frequent checkout routines; in many cases critical items are turned on continuously and monitored, an instance of infinitely frequent testing as a means of dealing with a near zero-time reaction goal.

Spare Electronic Boxes

In many cases electronic boxes are available for integration into a system long before the system build-up process is ready for them. The boxes are then stored until they are needed. In a normal assembly situation the questions of test frequency in storage are not acute since the time of use is hopefully known well in advance. The situation is more critical in the case of a spare box required only in the event of a failure of a primary unit already integrated into the system. Such a spare box is indeed subject to a random demand: its rapid availability to the system is more critical than normal since the spare is only required when some other anomaly has already occurred and schedules are very likely in jeopardy. Finding that the spare has also failed when it is taken from storage would certainly insure (i) significant delays; (ii) imposition of penalties for tardy delivery if such provision were in the contract; and (iii) a damage to the company reputation. Stored spares could be treated exactly as stored spacecraft were in the examples above.

Spare Tire in a Passenger Car

The previous example hits particularly close to home if you have ever had a flat tire only to find that your spare is also flat when it is pulled from its hiding place. How often do you check the air pressure in your spare?

Learning Theory in Psychology

Knowledge is stored in the brain and called upon in special situations. The longer it has been since some class of knowledge has been called upon, the less is ones assurance that it will be accessible when needed. Reviews of previously learned information correspond to the storage tests of our model; forgotten data corresponds to defects uncovered by storage tests; relearning forgotten data is the repair action adopted to remedy the failure. An interesting learning theory study could be based upon this storage model, investigating the optimum review frequency and intensity for various types of information, so that an adequate recall would occur when the data were needed at some unexpected time.

Relief Pitcher in Baseball

The relief pitchers lolling in the bullpen are actually in storage awaiting an unscheduled demand for their services. The more days which go by without his pitching, the less becomes the manager's confidence that the pitcher will "have it" when called up with the bases loaded. The lower the manager's confidence, the less frequently does he call upon the pitcher. This cycle leads to the often observed dependence of a manager upon one or two relief pitchers whom he tests often enough to have confidence in. They each appear in roughly half of the games in a season and typically burn themselves out in two or three seasons. A manager wishing to profit from aerospace technology would evaluate existing data on times-between appearances and its correlation with performance for each of his relief pitchers.

An analysis using a framework similar to that found in the storage model would doubtless identify a maximum number of days off which each pitcher can typically tolerate. Periodic games pitched entirely by little used relief pitchers would be one way to keep his entire staff in shape to be called upon when needed. This approach is also analogous to peridic calibration of electronic test equipment. No pitcher should be allowed to exceed his recommended calibration periods.

REFERENCES

1. Nathan, I., "Systems Availability Modeling, Considering Complex and Imperfect Sparing," presented at 1966 Aerospace Conference, Seattle, Washington.

2. Cottrell, D.F., Gangnier, T.R., and Kimball, E.W., "Dormant Operating and Storage Effects on Electronic and Part Reliability," RADC-TR-67-307, Martin Mariette Corporation, July 1967.

3. Church, H.F., and Roberts, R.C., "Failure Mechanisms of Electronic Components," The Electrical Research Association, Leatherhead, Surrey, England.

4. Neuner, G.E., "Research on the Reliability of Electronic Parts in Storage," TRW Report No. D00461, 72-2286.57, April 1972.

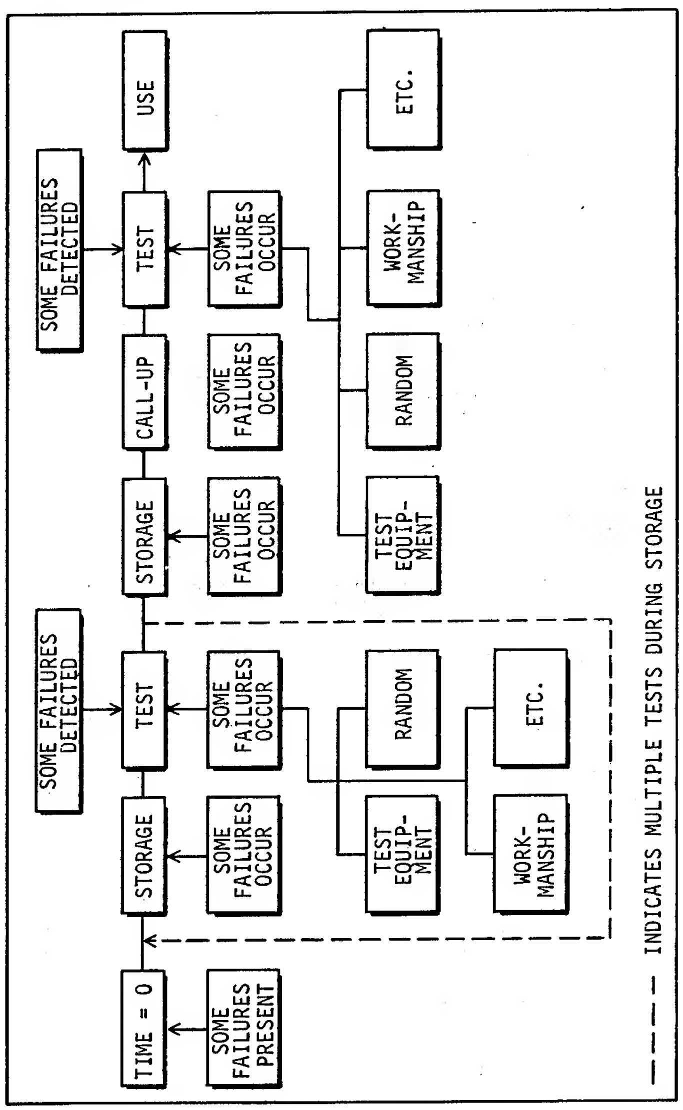


Figure 1. Flow of Storage Events

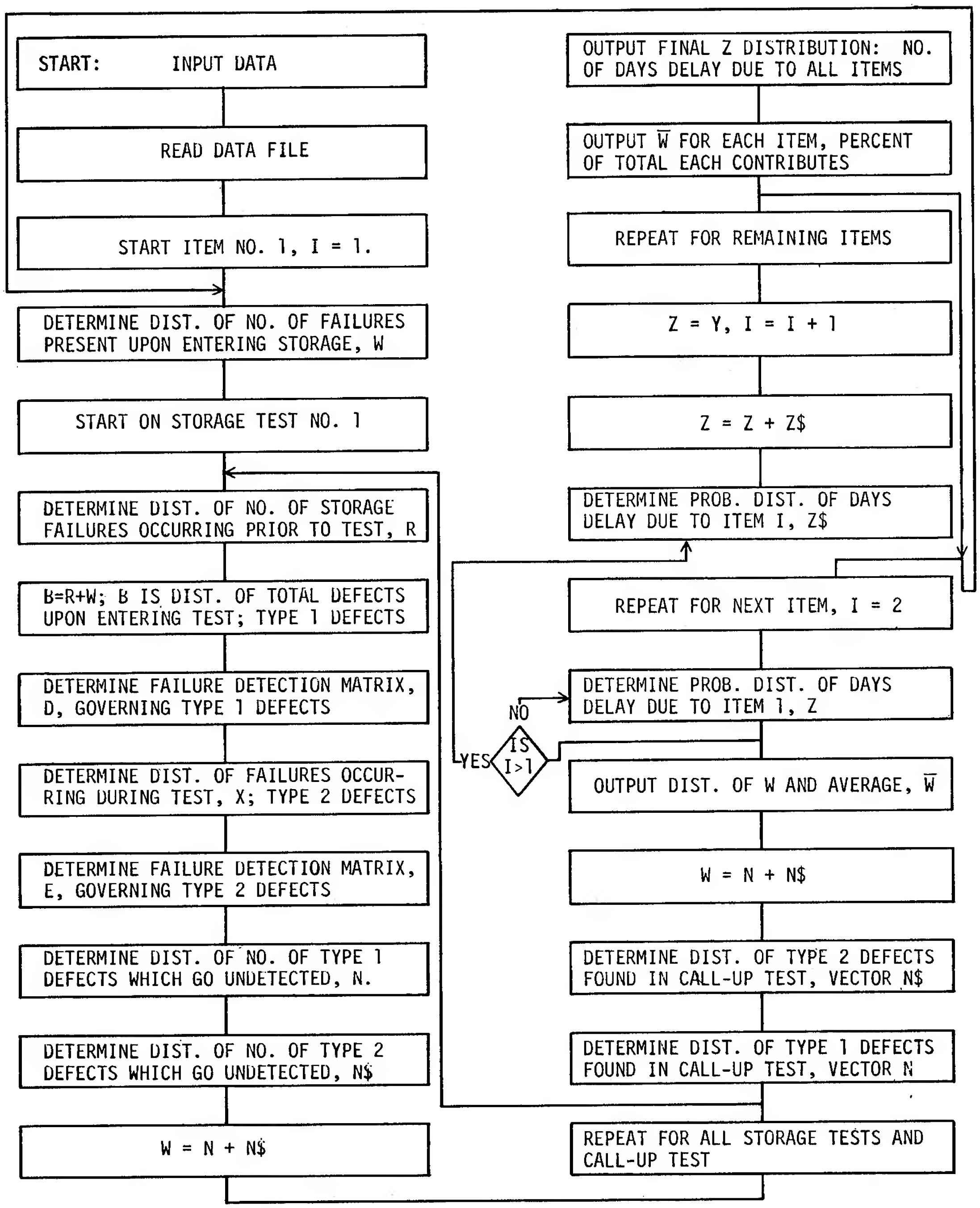


Figure 2. Computer Program Flow Diagram

MAGNETIC-TAPE QUALIFICATION AND ACCEPTANCE TESTING

by
T. J. Dylewski & R. H. Spitler
Information Processing (19-01, 102)
Lockheed Missiles & Space Co., Inc.
Sunnyvale, Ca., 94088

INDEX SERIAL NUMBER – 1060

I – SUMMARY

Twenty reels of each of eleven brands of magnetic recording tape have been tested for major and minor errors. This information was then used for identifying acceptable brands of tape and for developing incoming inspection acceptance procedures.

The sample size of twenty is regarded as the smallest that can be used without seriously degrading the value of inferences from test data. The average number of major errors per reel was taken as the measure of tape quality, since it is more suitable for this purpose than the proportion of reels without major errors.

The following conclusions have been reached:

1. Major errors follow the Poisson distribution.
2. Major and minor errors are positively correlated.
3. The results of this tape study and the preceding one are equivalent, although different testing equipment was employed.
4. The set of eleven brands tested can be partitioned into three homogeneous subsets according to quality.
5. Only brands in the highest quality subset are acceptable for LMSC digital-computer use.
6. Two brands of the nine brands that had also been tested in an earlier fourth tape study showed extreme changes in quality. The quality of the other seven brands did not change significantly.
7. Yearly repetition of the tape study is desirable in order to assure having up-to-date tape-quality information.

II – INTRODUCTION

From time to time a study is undertaken at Lockheed Missiles and Space Company, Inc. of various brands of magnetic recording tape for computer use. The results of the studies identify brands deemed acceptable and provide the basis for incoming-inspection rules.

Three earlier studies[1,2,3] employed the sophisticated statistical technique of analysis of variance in order to isolate effects of various hypothesized causes of tape performance variation. This is a sound approach in a new field with limited and highly variable data. By means of this technique it was possible to demonstrate that generally the principal cause of variation was the brand. Moreover, when other causes assumed some significance, their presence was usually independently apparent during or after the testing, and could be allowed for in the interpretation of test results. The fourth tape study[4] had the same ultimate objectives as the preceding ones; it concerned itself, however, solely with performance variation between brands. This is also true of the present, fifth study.

The plan[5] for the study outlined the procedures for the test phase and the analysis phase. Both phases have been carried out substantially as planned, so that the purpose of the study has been completely realized. Although unforeseen contingencies have caused some deviations from the plan, the validity and utility of desired study results have not been adversely affected.

The testing of magnetic recording tape is neither an accurate nor a precise discipline. The definition of what constitutes a defect is rather arbitrary, as is the allowable frequency of occurrence of defects in acceptable tape.

Tape evaluators used for testing magnetic recording tape attempt to record on magnetic tape and to read the recording in a manner which will disclose defects in the tape. It is expected that the tape defects found by the evaluator are approximately the same ones that would affect computer operation. Now, successful recording and reading of tape is the consequence of many interrelated discrete tape and equipment characteristics. Since we can observe only the end result of this involved process, a tape defect can be defined only in terms of performance under well specified conditions.

Unfortunately, nominal operating specifications affecting tape performance vary for different models of computers and tape evaluators. Thus, no single test procedure can claim to yield universally valid results. Moreover, test results are not always reproducible since tape and computer or evaluator performance is strongly influenced by cleanliness, drifting of equipment characteristics, and frequency and care of recalibration. It is well known that head wear, accumulation of iron oxide, and variation with age of vacuum and electronics can seriously affect observed tape performance. In addition to these problems, IBM has stated that its calibration tapes can vary ±5%. Of course, if they are duplicated as is done to reduce cost, then the variation could be considerably larger.

These circumstances make absolute quantitative rating of magnetic tape extremely difficult if not impossible, since variations of results by orders of magnitude can occur as a consequence of differences in testing techniques and conditions. Nevertheless, it is possible to utilize test results in a relative way for identifying acceptable and unacceptable tapes. The validity of such relative ratings depends upon meticulous attention to random sampling and adherence to a simple but fixed testing procedures in order to ensure that precision-reducing effects influence all tests in an unbiased and statistically identical manner.

III – TEST PHASE

1. Test Material

Twenty reels of digital-computer magnetic tape were purchased from eleven sources, identified by letters A ---K.

The purchase orders were worded as follows: "Tape, magnetic, digital, computer, certified 1600 BPI (3200 FCI), 1/2" × 2400', long wear, heavy-duty Mylar, total-surface tested, with 1108 leaders, solid flange aluminum hub reel for use on UNIVAC, IBM, Telex compatible tape drives, bulk pack, for test and evaluation."

As the test material was received it was stored together under normal room conditions and not used.

2. Test Procedure

a. Testing was conducted June 30 to July, 1971 at LMSC by properly qualified and supervised operators.
b. Each reel within a brand was sequentially numbered. The reels were chosen for testing in a sequence determined by a table of random numbers.
c. All testing was done on the same Data Devices cleaner-evaluator, Model 7900, identified as evaluator A. The machine was operated in the clean-test mode, full cycle. The playback level was set at 22% for major errors and 35% for minor errors. An error was determined to exist when "ones" written in each of 16 track positions were not read

back correctly, as indicated by a parity error. The number of major errors and the sum of major and minor errors were recorded on separate counters. During the testing for errors, the cylinder cleaning blade and brush were removed and only the read-write head was used. The head did not contact the tape until the rewind mode.

d. All capstans, roller guides, and heads were cleaned after each reel of tape had been tested.

e. The evaluator was calibrated before testing started and, on the average, after every 10 reels tested.

3. Test Data

Table I shows the observed frequency of occurrence of major and minor errors.

IV – ANALYSIS PHASE

1. Theoretical Considerations

The occurrence and non-occurrence of error-free reels in a sample of reels is a binary process correctly described by the binomial distribution. Still, since high reproducibility of error tests cannot be expected, the simple binary classification of tapes as good or bad is too tenuous to support acceptance and rejection decisions. Moreover, as a basis for analysis of tape test data, the binomial distribution is inefficient in the utilization of available information. Tape tests provide data on the frequency of occurrence of errors, but the binary approach disregards this fact by categorizing all tapes simply as either good or bad. Finally, the principal purpose of tape testing is to minimize the tape user's risk of loss, and that risk is much more highly dependent on the average number of errors per reel rather than on the simple absence or presence of error.

Thus, although the proportion of error-free tapes is a valid statistic, it does not possess as high a utility as the error rate. It is judged that the average number of major errors per reel is the most suitable criterion of acceptability.

If the probability of occurrence of an error in a given length of tape is constant, its frequency of occurrence is correctly described by the Poisson distribution. If all reels of tape are of standard length, it is convenient to state the rate of error occurrence in terms of errors per reel. Calculation of this error rate is done by adding the number of errors in all reels of one sample and dividing by the sample size. The Chi-square test is suitable for estimating how well the error-rate data can be described by the Poisson distribution; however, as this test becomes unreliable when the number of items per class decreases below 5, the Kolmogorov-Smirnov test is preferred[6]. Confidence limits for estimates of population error-rates can be taken from tables or graphs of the cumulative Poisson distribution[7].

If the observed error-rate data conform closely to the Poisson distribution, considerable confidence can be placed in the plausibility of these conjectures:

a. The testing technique was adequate.

b. The vendor meets performance specifications by controlling the manufacturing process so as to prevent the producing of bad tapes, rather than by eliminating bad tapes resulting from lack of process control.

c. The best tapes produced are not skimmed off for sale at a premium.

d. The tapes came from a single production process.

e. Tapes purchased in the future may be expected to exhibit similar uniform performance.

Conversely, if the observed occurrence of errors does not follow the Poisson distribution closely, one or more of the foregoing conjectores may be false.

Both major and minor errors are deleterious to computer operations, but experience indicates that the usually more numerous minor errors are also less reproducible under test. It is desirable, therefore, to investigate what relation, if any, exists between the two kinds of errors, with a view to obviating the use of minor-error data for setting acceptability criteria. For investigating this relation, Pearson's product-moment correlation coefficient is deemed adequate, since no suitable basis is available for hypothesizing casual or functional relations.

Neither is there an adequate before-the-fact basis for establishing a definite numerical boundary for dividing the tested brands into acceptable and unacceptable subsets. The boundary will be determined after-the-fact by partitioning the set of all brands into two or more subsets in such a way that there will be maximum homogeneity within each subset and greatest disparity between them. In order to achieve this end, that partitioning will be adopted which leads to a minimum value of the sum of squares of deviations of individual brand error rates about respective subset error-rate

TABLE I

OBSERVED OCCURRENCE OF ERRORS IN 20 REELS
(Maj = Major Errors; Min = Minor Errors)

Errors Per Reel	TAPE SOURCE																					
	A		B		C		D		E		F		G		H		I		J		K	
	Maj	Min	Maj	Min	Maj	Min	Maj	Min	Maj	Min	Maj	Min	Maj	Min	Maj	Min	Maj	Min	Maj	Min	Maj	Min
0	15	7	12	7	8		8	2	9	1	3	1	14	4	1				4		1	
1	4	8	4	4	9	4	9	3	7	8	4		1	5	1	1	2		4			
2	1	2	2	3	3	5	2	6	2	6	3	2	3	7	5				4	1		1
3-4		3	2	4		5	1	7	2	2	6	4		2	5	2	3	1	4	2	5	
5-8				2		6		2		2	4	9	1	1	7	2	11	2	3	5	8	
9-16												4				6	3	5		5	4	9
17-32										1						3		6		5		7
33-64													1		1	5	1	5		1	1	1
65-128														1		1		1				1
129-256																						
257-512																			1			
513-1024																					1	1
1025-2048																				1		
Av. Maj.	.300		.750		.750		.800		.850		2.65		2.70		6.05		7.45		15.45		34.15	
Av. Min.	1.05		1.80		3.45		2.60		3.35		6.40		7.90		21.65		29.40		75.50		64.05	

means. This method of partitioning a population into homogeneous subpopulations has been developed by T. J. Dylewski and has been employed on various occasions for forming homogeneous groups.

Since the nine brands tested in the preceding study were also tested during the present one, the results of the two studies can be compared. The most suitable statistic for such a comparison is a rank-correlation coefficient, such as Kendall's[8], because it permits inclusion of data on three brands for which it was not possible to measure error rate for some reels of badly damaged tape in the fourth study.

High positive correlation between the results of the two studies would support conjectures that:

a. The relative quality of the various brands has not changed much with time.
b. The suppliers have furnished representative samples.
c. The two testing procedures were equivalent.
d. The testing procedures were adequate.

Conversely, lack of high positive correlation would mean that one or more of these conjectures may be false.

2. Poisson Confidence Limits

For an ideal Poisson distribution, the average and the variance have exactly the same value. If two or more different Poisson distributions are merged, the variance becomes greater than the average; this was observed in nine cases. If the upper end of the series is eliminated, the variance becomes smaller than the average; this was observed in two cases, but the effect was not large.

The Kolmogorov-Smirnov test was employed for testing the hypothesis that error rate follows the Poisson distribution. Table II presents the results of this test for all eleven brands.

TABLE II

Kolmogorov-Smirnov Test of Hypothesis That Major-Error Rate Follows the Poisson Distribution

Tape Source	Observed Average Major-Error Rate, in Errors Per Reel	Variance of Observed Error Rate	Maximum Relative Difference (D) Between Poisson and Observed Distribution	Probability That a Greater Value of D Could Have Occurred by Chance
A	0.300	0.478	0.013	>0.20
B	0.750	1.355	0.127	>0.20
C	0.750	0.518	0.073	>0.20
D	0.800	0.689	0.049	>0.20
E	0.850	0.898	0.045	>0.20
F	2.650	4.029	0.097	>0.20
G	2.70	87.17	0.356	<0.01
H	6.05	102.0	0.322	0.05−0.01
I	7.45	53.10	0.253	0.15−0.10
J	15.45	3537.	0.944	<0.01
K	34.15	13,630.	0.900	<0.01

For the six brands having the lowest error rates, the observed distribution of major errors does not differ significantly from an expected Poisson distribution. For two of the other brands the difference may be significant; and for the remaining three it is definitely significant. In particular, major errors in the G, J, and K tapes cannot be regarded as following the Poisson distribution. The case is marginal for H and I.

Table III presents estimates of population major-error-rate limits at the 0.95 and 0.99 confidence levels for brands whose observed errors definitely follow the Poisson distribution.

TABLE III

0.95 and 0.99 Confidence Limits Based on Poisson Distribution of Major Errors

Tape Source	Observed Average Major-Error-Rate, in Errors Per Reel	At the Stated Confidence Level (P), the Sample Came From a Population in Which the Major-Error Rate Can Have This Range: P = 0.95	P = 0.99
A	0.300	0.008−4.254	0.002−5.939
B	0.750	0.018−5.100	0.004−6.898
C	0.750	0.018−5.100	0.004−6.898
D	0.800	0.020−5.194	0.004−7.004
E	0.850	0.022−5.288	0.004−7.110
F	2.650	0.487−8.228	0.256−10.382

3. Correlation Between Major and Minor Errors

Pearson's product-moment correlation coefficient for major and minor errors in all eleven brands is 0.862. This value or a higher one could arise from random sampling of an uncorrelated population less than 0.001 of the time.

4. Homogeneous Quality Grouping

The least-squares partitioning procedure was applied to data for major errors. These three homogeneous quality groups were obtained:

HIGHEST QUALITY	INTERMEDIATE QUALITY	LOWEST QUALITY
A	F	H
B	G	I
C		J
D		K
E		

5. Correlation Between Results of Fourth and Fifth Studies

Table IV gives the rank of the nine brands tested during the present study and the preceding one. Ranking was employed here because in the previous study D, H, and J samples contained reels of tape in such poor conditions that they recorded errors continuously. Thus, these three brands were assigned ranks based on the number of reels of undamaged tape, whereas the other brands were ranked according to the number of permanent errors per reel. In the present study all reels were undamaged and the brands were ranked according to the number of major errors per reel.

Kendall's rank-correlation coefficient for the pairs of data for all nine brands is 0.22. This value or a larger one could have arisen by chance 0.53 of the time as a consequence of random sampling of an uncorrelated population.

TABLE IV

Comparison of Results of Fourth and Fifth Studies

Tape Source	Rank of All 9 Brands Common to Both Studies		Rank of 7 Brands, Common to Both Studies, Whose Rank Changed Least	
	Rank in Fourth Study	Rank in Fifth Study	Rank in Fourth Study	Rank in Fifth Study
A	4	1	3	1
C	2	2	1	2
D	7	3		
E	5	4	4	3
G	3	5	2	4
H	8	6	6	5
I	1	7		
J	9	8	7	6
K	6	9	5	7

It is evident, however, that only brands D and I have experienced a large change in rank. If the other seven brands are considered alone, Kendall's rank-correlation coefficient for the results of the two studies becomes 0.57. This value or a larger one could arise by change 0.095 of the time as a consequence of random sampling of an uncorrelated population.

V – CONCLUSIONS AND RECOMMENDATIONS

The present study has provided additional insights regarding magnetic-tape error-statistics, and their estimation and interpretation.

The rate of occurrence of major errors in the six highest-ranked brands follows the Poisson distribution. Accordingly, the theoretical Poisson distribution can be employed validly for estimating confidence limits from the observed data. It can be reliably inferred that the Poisson distribution would be perfectly demonstrated under ideal tape production and testing conditions. Thus, the way in which observed error frequencies depart from the Poisson distribution can be informative about process quality control and testing techniques.

The randomization, cleaning, and calibration procedures during the test phase of the study were planned and executed so as to provide statistically identical testing conditions for all tape brands. Thus, failure of major errors in the lowest-ranked brands to conform to the Poisson distribution must be taken as a sign of poor quality control at the source. Furthermore, since the variance is much higher than the average error rate for the low-quality brands, it can be inferred that the samples are inhomogeneous groups containing items from more than one process, and may include outputs of uncontrolled processes as well as customers' rejects.

The cumulative Poisson distribution has provided estimates of population major-error-rate limits which are in reasonable agreement with observed results. These confidence limits are a suitable basis for incoming-inspection specifications.

That major and minor errors might be positively correlated for some brand of tape is not an unreasonable supposition. A high positive and significant correlation has been demonstrated on a much broader basis by merging the data for all the samples. Accordingly, investigation of one type of error should usually suffice for tape evaluation.

Of the eleven brands of tape tested, only A, B, C, D, and E are deemed acceptable for LMSC digital-computer use. This group is not divisible further by the homogeneous-grouping technique on the basis of average numbers of major errors. It should be noted that A, E, D, and C exhibit, in that order, best conformance to the Poisson distribution and can be supposed to come from more closely controlled processes than the rest of the brands. Since for D and C the variance of the error rate is lower than its average, it can also be supposed that these sources may be employing procedures for weeding out low-quality tapes.

Failure to find significant correlation between the results of the fourth and the fifth tape studies for the nine brands common to both studies is ascribable to two possibilities. The testing procedures of the two studies may not have been equivalent, and the quality of the tapes furnished by suppliers may have varied with time.

Test procedures were indeed different in the two studies. Formerly, a Cybetronics certifier, Model 1600, had been used for identifying permanent and temporary errors. Now, a Data-Devices cleaner-evaluator, Model 7900, has been used for identifying major and minor errors. In the earlier instance, a cleaning operation was performed in an attempt to remove a detected error; if the error could be removed, it was called temporary – otherwise, it was recorded as permanent. In the latest study, difference in playback level served to distinguish between major and minor errors; error removal was not attempted.

From the foregoing it can be expected that, even for the same reel of tape, every case of permanent error is not necessarily a case of major error, and conversely. Still, the two test procedures cannot be arbitrarily regarded as so different as to produce completely uncorrelated results, because extreme differences in rank occurred only in the case of the D and I brands. When these two brands are removed from consideration, the results of the two studies for the other seven brands are so much more strongly correlated that they may be taken as equivalent. Most of the small disparity between them can be reasonably assumed to be the consequence of changes in within-brand tape quality during the two years separating the two studies. Strong support for this point of view resides in the fact that almost 20% of the tape furnished for the fourth study was so badly damaged that it could not be tested in the normal manner.

The distinct possibility that the quality of supplied tape may change markedly over time prompts the recommendation that the tape study be repeated yearly in the interest of keeping information on tape quality up-to-date.

The results of this study form a suitable basis for inspection plans in accordance with military inspection practices[9].

VI – BIBLIOGRAPHY

1. H. P. Hartkemeir, A Statistical Study of the Performance of Four Brands of Magnetic Tapes and The Channels and Tape Drives of the IBM 7090, LMSD-895042, Feb. 1961.
2. H. P. Hartkemeier, The Performance of Information Processing Equipment, or an Industrial Application of Analysis of Variance to Performance of Three Brands of Long Wear Magnetic Tape and Two Channels and Four Tape Drives of the IBM 7090, 5-11-61-1, Oct. 1961.
3. R. H. Spitler and H. P. Hartkemeier, Statistical Analysis of Performance Data on Evaluation of Eight Brands of Magnetic Tape, LMSC-678864, Oct. 1967.
4. T. J. Dylewski and R. H. Spitler, Report on Fourth Study of Magnetic Tape, LMSC-690576, Aug. 1969.
5. T. J. Dylewski and R. H. Spitler, Plan for Fifth Tape Study, TN 14-1, March 1971.
6. S. Siegel, Nonparametric Statistics, McGraw-Hill, N.Y. 1956, pp 47–52.
7. H. F. Dodge and H. G. Romig, Sampling Inspection Tables, John Wiley, N.Y. 1959.
8. M. G. Kendall, Rank Correlation Methods, Hafner Publishing Company, New York, 1955.
9. -----, MIL-STD-105D, Sampling Procedures and Tables for Inspection by Attributes, Department of Defense, Washington, D. C., 1963.

Product Liability Prevention Via A Controlled System

Alexander Sternberg
Systems Consultant
Trenton, New Jersey

INDEX SERIAL NUMBER – 1061

Abstract: Product liability has become a major factor impacting American industry. Billions of dollars are being paid out annually to resolve product and service liability claims instituted by today's consumers. The philosophy of "Do It Right The First Time" becomes more important than it ever was before. Many companies have had to go into bankruptcy after being attacked in a product liability suit. The need for an approach to assist in preventing the dilemma pictured is urgent. This paper has integrated existing state-of-the art techniques into a Control System and presents one type approach which will both assist in resolving existing product liability problems and also minimize the occurrence of future problems using the techniques related to control.

Introduction:

The Age of Consumerism has resulted in very dynamic and significant changes in the marketplace. Since the early 1960's these changes have become most profound and cataclysmic. Legislatures have modified their interpretations of the law, the consumer has become more outspoken and critical, and consumer advocate groups have sprung up to assist the consumer in this battle against American industry. The need for communication across all involved disciplines has become a necessity for survival. When we further realize that the United States in 1971 has become the only community in the world wherein the gross national product has greater than 60% of its total dollars attributed to the service industries, the situation becomes even more emphasized. The problem becomes more complex when we observe that the systems and procedures used by the service industries are generally backward and antiquated and that the Industrial Revolution has not arrived as yet. The quality of the service performed and the management control exercized leaves much to be desired. This, unfortunately, is the most important factor causing the problems faced by American industry in today's marketplace. This presentation is intended to demonstrate a mechanism which can resolve this tragic and most costly issue. The writer contends that the efforts discussed in this paper can be implemented at minimal additional cost to the company, although there is need for a management organization to implement the prevention program consisting of representatives from the departments involved such as legal, insurance, engineering, manufacturing, advertising, packaging, quality control, etc. Most companies presently have personnel performing the tasks involved, however the effort is generally not integrated and uncoordinated. No one person has been given the authority or the responsibility for managing the product liability prevention program. Someone high enough in management who can cut across the many disciplines and tie together the complex package is required. Unfortunately the various individual motivational forces of the involved departments do much to retard this effort. It is essential that these forces be eliminated or overcome. The goals of the program have to be established as company goals with all groups striving toward these common objectives. Further the program should be costed and established as a key line on the profit and loss statement and measureable via the established accounting system. Unless this is accomplished it becomes another platitude with minimal meaning to all concerned. If management is going to invest in a program, there must be a measureable return on investment, otherwise it will be the first area considered for elimination when pressures are exerted for cost reduction. It is also important to include representatives of all departments concerned early in the planning of the program. Goals have to be established with the concurrence and participation of these concerned people.

History:

It is worthwhile at this point to present a picture of the product liability story for illustrative purposes. The story unfolds when we review what has taken place. In the past the environment is well depicted by the expression "Caveat Emptor" or Buyer Beware. With the advent of the Industrial Revolution, stress was placed on encouraging the growth of American industry. This was further reflected in the decisions made by the various courts whereby unless privity of contract existed fault was not attributable to the manufacturer. This picture changed with the decision by Justice Francis in the Henningsen versus Bloomfield Motors case in 1960 eliminating the need for privity of contractual agreement to prosecute a case of strict liability. Since this decision, major changes have taken place. The significance of these changes can best be illustrated by using the history of what has transpired in the growth of strict liability cases since 1960.

In 1960 in Cook County, Illinois, there were three decisions with a total aware of $4,211, an average of $1,400 per case. In 1966, in this same Cook County there were eight cases. Excluding one abnormal case of $725,000, the average award from the seven remaining cases was $33,666, an increase of 2,000% in six years. This picture is further emphasized when we review the growth of claims in the United States. In 1960, there were several hundred cases; 1963 saw 50,000 cases on the court dockets. This exceeded 100,000 in 1968. The number grew to 500,000 in 1970 and is expected to reach 1,000,000 by 1973. It is most significant to realize that most cases involve more than one defendant. All parties responsible for the design, manufacture, inspection, sales, and distribution of the product may become defendants in a product liability case. During the trial the lawyers may subpoena the president,

design engineer, quality control engineer, and all those considered to be directly or indirectly responsible for allowing defective product to be produced. The results of the last ten years are well illustrated by the figure below.

RESULT OF THE PAST TEN YEARS

1. Federal Auto Safety Standards
2. Federal Tire Standards
3. Truth-In Lending
4. Truth-In Advertising
5. OSHA 1970
6. Construction Equipment Safety Standards
7. Instruction Decals and Warning Labels
8. Action Line - Value Line
9. Consumer Protection Agencies
10. Consumer Affairs Groups
11. Ombudsmen
12. Radiation Hazards Act

Figure 1

Problem: It is important to define the problem itself before proceeding further. Figure 2 below summarizes this but it may be most appropos to discuss it more in detail.

THE PROBLEM

1. More aggressive and demanding customer
2. Higher standards of quality and service
3. Federal and state governmental intervention
4. Increases in claims and losses
5. Something for nothing philosophy
6. Organized plaintiffs' bar
7. Politicians looking for publicity
8. Newspapers seek the sensational
9. Reinterpretation of the law

Figure 2

We have before us a very dynamic type marketplace. The customer today is fairly well educated and informed and has become more aggressive and demanding. He senses his power and is not hesitant in exerting it. He has developed a sense of requiring higher standards of quality and service and, in turn, demands it. The federal and state governments are responsive to this and are intervening in his behalf via consumer protection legislation such as truth in advertising, lending, radiation hazards act, occupational safety and health act, and others. The public has developed a growing claims consciousness and is not hesitating to institute legal action. With the introduction of the "something for nothing" philosophy there is less reluctance to seek legal recourse if a question arises as to fault for injury.

The organization of the American Trial Lawyers Association has now set up an organized plaintiff's bar. The public has a defender to assist in this endeavor. This group works together providing specialists in liability litigation for the consumer. This situation is further enhanced by the fact that politicians are using the various issues involved as a means of publicizing themselves and seeking votes. The newspapers are also riding the bandwagon seeking the sensational via articles in this area increasing their circulation. The law is now being more liberally interpreted to favor the public. It is certainly not uncommon to see the development of such groups as Office of Consumer Affairs, Bureau of Consumer Protection, Action Line, Value Line, Mr. Fixit, etc. who are organized to assist the consumer in resolving his problems with American industry. Putting the picture together certainly paints a defensive pattern for industry.

Records: When we then examine the type of information most often required to prevent and defend against product liability litigation, the picture becomes more vivid. This list is summarized in the figure below.

RECORDS HELPFUL IN DEFENSE

1. Blueprints and schematics
2. Rejection reports; Acceptance reasons
3. Quality Control Procedure Checklists
4. Reject history
5. Quality Control Manual
6. Action taken on suggestions for reducing defects
7. Inspection and test procedure records
8. Laboratory test reports
9. Compliance with government regulations
10. Sales literature
11. Sales records
12. Sales slip showing warranty
13. Checklists covering inclusion of instruction manuals in shipment
14. Field failure reports
15. Feedback from salesmen
16. Past liability claims
17. Statements from witnesses
18. Photos before and after

Figure 3

The other cause of litigation is a result of negligence errors. The writer has summarized these causes in the figure below:

NEGLIGENCE TYPE ERRORS

1. True design errors
2. Inadequate safety devices
3. Use of failed safety devices
4. No post manufacture safety check
5. Use of unsafe or unsuitable material
6. Inaccurately planned manufacturing process
7. Lack of planning for foreseeable uses
8. Unforeseen consequences of wear and tear
9. Use of unnecessary part
10. Below industry standard level
11. Ignorance of scientific knowledge throughout industry
12. Inadequate warning or failure to warn

Figure 4

The Control System "Elements":

It is the opinion of the writer that a majority of the problems relative to product liability can be prevented by the use of a controlled system approach incorporating the implementation in a timely manner of guideline documents and quality and reliability methods which are discussed in this paper. These certainly are not novel but rather common sense procedures organized in a manner to provide timely information and a mechanism for establishing a closed loop feedback system. These prevention tools are the integral elements of a management information system. These items have been categorized into two sections, namely, guideline documents which are used as baseline procedures by all concerned company personnel and quality and reliability methods which are key elements of a management control system. A review of these, in detail, will explain their applicability to the control system concept discussed here.

A. Guideline Documents: (Engineering documentation)

In order for the product to be manufactured at a minimum cost and meeting the standards of producibility, certain information is essential. These are primarily engineering standards. For these standards to have mutual understanding for all personnel concerned, it is necessary to prepare them in accordance with standard procedures. These procedures are defined by the writer to be Guideline documents. These Guideline documents are prepared as guidance procedures for the use of personnel responsible for preparing engineering specifications and therefore establish the baseline to be utilized to assure that major criteria have not been overlooked in their preparation. A list of these documents and an explanation of how they should be used follows:

1. Workmanship standards
   These standards are generally prepared by the Quality Assurance function coordinated with engineering and manufacturing personnel. They include such items as acceptable standards for soldering, welding, burrs, nicks, surface finish, potting, encapsulation, electrical and mechanical connections, tolerance, parallelism, etc. Unless some special requirements are required, these automatically govern and establish the workmanship requirements for all items manufactured or purchased. It is through conformance to these standards that the workmanship level of the product manufactured is attained to whatever level established.

2. Specification content and format guidelines.
   Unfortunately, whether it be due to an engineers nature to want to be different or the desire to be inventive, there is a tendency existing to prepare specifications in as many ways as there may be people involved. This breeds confusion and much misunderstanding as well as less of clarity. It helps considerably to develop guidelines to assist engineers in defining what should be included within a specific specification as well as how the specification should be organized. This should include checklists which allow the originator of the specification to review major criteria which should be considered leaving the decision to him whether a specific criterion is applicable. This establishes a means for assuring that all major criteria have been considered during specification preparation.

3. Guidelines, When and How to use a specification
   The importance of a well defined and practical specification can not be underestimated. In this context it is most helpful to establish a guideline which defines when a specific specification is applicable and how it should be used. Again the stress on uniformity and the development of a specification system which is readily identifiable is most important. This sets up a situation whereby ready reference is available to the proper documentation for all people concerned. When an individual desires to locate some information he has a system installed which blends with this objective. He also knows specifically where he should place specific information for the use of others.

4. Drafting Standards
   It is so critical to be certain that what we make or buy is specified correctly and adequately. This is obviously most applicable to how a drawing is prepared and what must be specified on it. The delineation of satisfactory tolerances, concentricity, parallelism, dimension, angles, alignment of parts, edges, and the like govern the product which will result. The need for correctly defining the proper material and processes also drastically affects the product quality. All this can be most adequately covered in the development and preparation of Drafting Standards. The requirement for the use of these standards by both the engineers and the draftsmen assures control of the product made or purchased.

5. Index to Standard's files
   What benefit is there in having detailed standards developed and prepared if they can not be referred to and readily used by personnel who need them? It is difficult to set up a set of standards, including the various types of specifications, procedures, and the like, without this library becoming very detailed, numerous, and complex. This is particularly so when there is more than one product type being manufactured. As a result, the library must include volumes of in-

formation. Regardless of the volume of date, this information must be easily accessible to cognizant plant personnel. An index to the file simply prepared both by item number and item title with a cross reference is essential toward attaining the desired objective. This index should also include information available from sources outside of the plant. Whether it be other parts of the company, national standards, government files, or others, it should be readily locatable.

6. Computer Programs
Another vehicle used to collect and tabulate information is the computer. This tool has opened many avenues which were originally either too costly or too time consuming to accomplish. This is especially true in the preparation of bills of material which serve as tables of contents detailing all the parts, materials, processes, sub-assemblies, and assemblies which make up each product type. One program used is to list all the subparts of a product as assemblies down to the raw material used. This includes all the inspection and test specifications involved related to the specific assembly level. When this date is compiled in the computer, different types of outputs can be generated. These can consist of Bills of Material by product type, where used by product type, Assembly Bills of Material by product type, where used by assembly level, processing by assembly level, Travel Tickets, Inspection and Test flow diagrams, costs by assembly level and many other variations. The achievable results are purely a function of the computer program developed and the subsequent data fed into the computer. The potential is fantastic. It can be another tool for the plant personnel to use.

7. Specifications Change Control Procedure
A specification system, regardless of how good it is serves no future purpose unless changes to it are controlled and documented. Unless we are able to determine and control the changes which have been made to the initial design and manufacturing process, we will never be able to assure ourselves that product quality, life and performance are not being degraded. If we do not know where we were and can not determine what we are doing, how can we know where we are going and whether the decision made was proper.

8. Engineering Test Procedure
It is too simple to run a test and draw erroneous conclusions. I recall one instance during my past work experience whereby an engineer conducted a test in the factory. He came out with fantastically successful results and felt like he had conquered the unconquerable. Unfortunately when his results were carefully reviewed it was noticed that he was refuting Ohm's Law. A repeat of the test contradicted his results and clearly indicated that he had conducted an uncontrolled test wherein many variables were interacting without any control being taken to assure the elimination of any biasness.

This only illustrates the importance of setting up a test properly and assuring that it is performed under controlled conditions. This is not the usual in most troubleshooting factory operations. The preparation of a test procedure defining the use of designed experiments and how a test should be run can be extremely useful to all personnel responsible for this type of effort.

B. Quality and Reliability Methodology
In addition to the engineering documentation baselines for management information, the quality and reliability methods play a significant role in this picture. Much has been discussed about these procedures in many other books and manuals. These are summarized below:

QUALITY AND RELIABILITY METHODS

Statistical Techniques - Design of experiments, control charts, sampling, analysis of means.

Inspection and Test Plans - Plans covering what, when, how, who inspects and/or tests.

Quality System and Product Audits - Audit system being used and product whether conforms to requirements.

Process Control - Quality of manufactured product meeting needs and controlling it before too late or too costly.

Test Surveillance - Tests being performed properly, etc.

Design Review - Review of design, starting with concept, and in greater depth as design firms reviewing status of effort resolving problems as they arise.

Contract Review - Needs of customers considered, is process capable of meeting it, etc.

Failure Reporting and Analysis - Closed loop feedback system

Source Inspection - Critical parts inspected at source to prevent problems later.

Purchase Order Review - Imposing proper requirements on vendors, shared liability risk, etc.

Specification Review - Adequacy, completeness, etc.

Tool & Equipment Calibration - Frequency, which have to be measured, what precision, etc.

Vendor Surveys - Select vendor with capability to do job.

Quality Cost Control - Isolate major quality cost areas and emphasize cause and effect studies in these areas.

Reliability Analysis - Stress analysis environmental studies, reliable parts, block diagram, circuit analysis, degradation analysis, worst case, etc.

Part Selection & Application - Reliable part used properly.

Human Factors Engineering - Safety laboratory, misuse of production considerations, design to prevent safety problems, etc.

Failure, Mode, Effects, Hazards & Criticality Analysis
Analysis of failure modes considering hazards, effects and criticality and evaluating for compensating provisions.

Reliability Prediction - part failure rates and circuit analysis to predict reliability

Process Capability Studies - Can process meet specifications, what controls needed, etc.

Fault Tree Analysis - analyzing for faults relative to elements of equipment and determining prevention means.

The Control System:
Since we have discussed the guideline documents and quality and reliability methods, it should be most appropriate to see how these fit into the control system philosophy.

A simple illustration of the control systems approach is demonstrated by an explanation of the activities associated with the flight of a commercial airplane. Figure 5 depicts the steps in the process from initial planning through the final step, the filing of the flight log. The total picture represents a simplified explanation of a control system. Figure 5 has been categorized into four sections, namely 5-1 through 5-4 depicting the elements of the simplified controlled system and is attached to this paper.

5-1 shows the activities associated with the initial planning for an airplane trip. This involves the consideration of weather, distance, load, fuel needs, safety factor, traffic navigation aids, aircraft checkout, and the filing of the flight plan.

5-2 delves into the area of take off and initial flight. Data feedback, prevention measures related to a storm anticipated ahead, communication and coordination with the communications centers on the ground and in the air, and a decision to commit to a change in the initial flight plan to bypass the storm clouds are all depicted.

5-3 shows the activities taking place during flight. It depicts the feedback and analysis involved and the corrective action taken to return to the original flight path after the storm has been evaded. 5-4 is the landing and flight completion. It is at this time that the flight log is filed and becomes the basis for historical information and use for future planning via feedback and implementation of the knowledge gained.

In total we have here a simplified Control System which is certainly similar to the program requirements of a Quality Control System. It can be directly related to ASQC standard C-1 illustrated in Figure 6 titled Quality System Program Requirements.

Quality System Program Requirements

1. Planning, direction, control
2. Responsibility assigned to authority
3. Technically competent staff
4. Sufficient authority
5. Written procedures
6. Information available
7. Task definitions maintained
8. Changes controlled
9. Control over purchases
10. Records
11. Corrective Action program

(ASQC Std C-1)

Figure 6

Taking the elements of the Quality System we see that here, again, is required adequate planning, forceful direction, and control in measurement, and evaluation of the effectiveness of the control system as was covered in the initial planning relative to the airplane flight.

Administration of the controlled system vested in a responsible authoritative element of the organization with clear access to management is a prerequisite of any management system. The system has to be staffed by technically competent personnel with freedom to make decisions unbiasly. There has to be sufficient authority and written quality control, test, and inspection procedures used, kept current, and maintained. Information has to be available and maintained to insure that the job is performed properly and standardized. The reminder of the requirements are adequately summarized in Figure 6. The point to consider is that we are discussing a control system serving as a tool to management for the effective approach toward preventing product liability problems.

Benefits:
Such a control system has many benefits. These can be readily summarized by the following items:

1. It leads toward a systematized approach
2. Emphasizes the development and implementation of standards
3. Results in increased productivity.
4. Minimum risk decisions by seeking out the key areas for investigation
5. Makes management aware of problems on a timely basis and leads to problem solving techniques.
6. Assures that the manufacture or the service performed is of a quality nature.
7. Establishes an improved competitive position for the company
8. Leads to reduced insurance rates for liability
9. Decreases costs of operation
10. Serves as a positive response to con-

sumer criticism.
11. Leads to reduced set up times before going ahead with any production work.

If we compare the elements described above to the items listed in a Product Loss Control Program it becomes obvious that the insurance industry is proposing a very similar approach toward product liability prevention as has already been stated. Let me take the liberty to list some of the key elements of a product loss control program as extracted from literature published by the insurance industry. These are as follows:

1. Develop a management philosophy to frankly and aggressively support product reliability in all phases of the business.
2. Set up a program to get everyone in the company into the act
3. Establish a continuous and firm line of communication between all personnel.
4. Top management should select some person at a high level of authority to direct the program
5. Maintain a supplier control program
6. Have a program to control non conforming materials
7. Design for safety
8. Have a continuing quality control program
9. There should be a new product comprehensive test program tied in with quality control
10. Establish clear records and accurate record keeping programs in house and outside.

When we compare the quality control program requirements to the product loss control program requirements, the similarity is obvious.

Conclusion:
It is the writer's intention to illustrate in the presentation how the elements of the control system may prevent product liability litigation by citing case histories and relating them to the tool which chould have prevented the problem encountered. Let me conclude by stressing what management needs to know.

This is the following:

1. Professional talent is required
2. Have a documented and functional system
3. Don't fight the system - use it.
4. There is a need for a total coordinated product liability prevention team.
5. All people must be trained and motivated.
6. There is a need for workmanship standards.
7. Make it like the blueprint
8. Management always retains the responsibility for the quality of the product.

It is most appropos to complete this presentation with a bit of philosophy considered extremely important to follow in order to successfully implement a product liability prevention program. The effort has to be established as a key task and therefore a line item has to be included in the budget. The results of the program must be measureable via the company's accounting system and show up as an item in the profit and loss statement. Savings must be determined as well as losses and comparisons between past and present made. It is also essential that all concerned personnel be integrally involved including the mutual establishment of goals, budgets, and schedules. These personnel must also continue to participate throughout the program and work together toward mutually established objectives. Unless this is done, there is no meaning to the program.

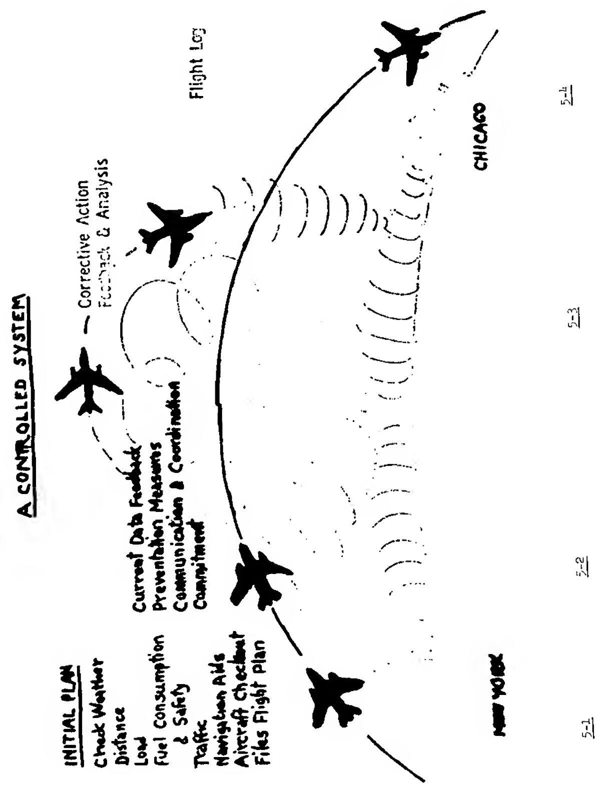
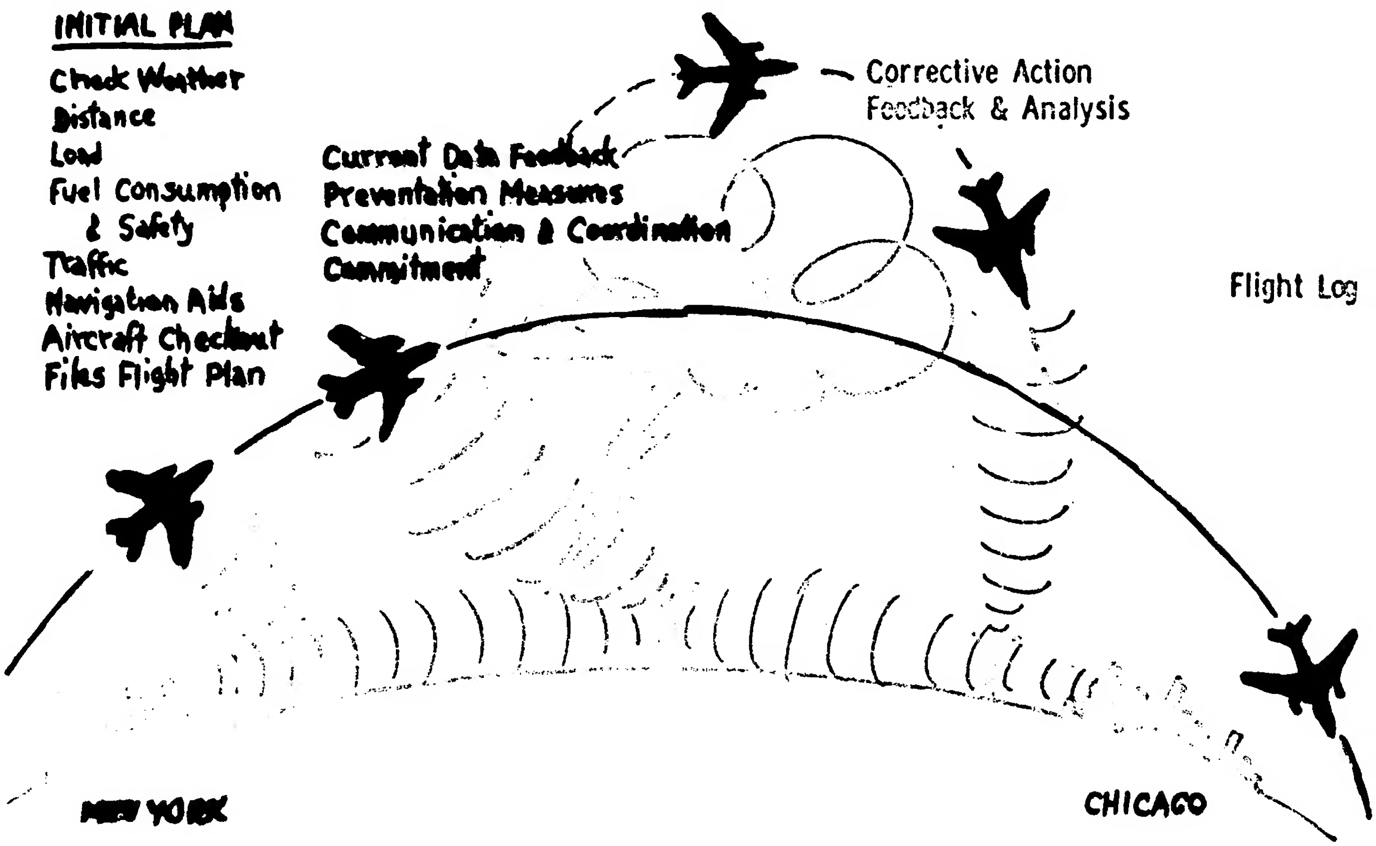


FIGURE 5

A CONTROLLED SYSTEM



5-1 5-2 5-3 5-4

FIGURE 5

J. R. Fragola and J. F. Spahn
Grumman Aerospace Corporation
Bethpage, New York 11714

Abstract

With the planned use of the Space Shuttle to deploy, repair, refurbish, and retrieve satellites, the models normally used for estimating satellite system reliability become unacceptable. The new capabilities for in-orbit replacement of satellite modules and reuse of satellites after earth-based refurbishment add several new dimensions to the design of satellite systems. In particular, performance for many systems would now be specified in terms of system uptime, not satellite life.

To deal with this increased complexity, two mathematical techniques, Markov modeling and dynamic programming, have been combined. Candidate systems are first modeled, a "best" configuration chosen, and then redundancy is allocated within an individual satellite.

As a practical test of this technique, the analysis was applied to the Large Space Telescope, a prime shuttle astronomy payload. The results indicate that the optimal system consists of two satellites, each with a 0.8 reliability for one year. More importantly, this method of analysis appears to be readily applicable to many future satellite systems.

Introduction

Satellite system economics are affected strongly by the advent of the Space Shuttle. A figure of merit such as satellite reliability now has meaning only in terms of the added flexibility gained by resupply. No longer can the inherent life of the satellite in a given system be considered equivalent to the system life. Problems which heretofore were solved simply now become more complex if full economic advantage is taken of the repair capability imparted by the shuttle. (See Fig. 1.)

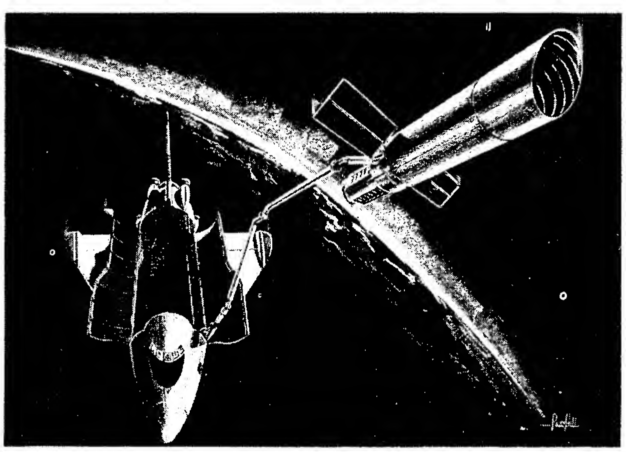


Fig. 1 Shuttle Manipulator Retrieving LST

In the past, the method of achieving program goals made the economics simple - the product of the number of satellites and the probability of success of each should be made maximum within the overall program cost constraint. Selection of these two variables was driven by direct and implied costs. For example, direct cost was the product of the unit cost to build a satellite with a given life and the number built. The indirect cost was the cost added to minimize the risk. Risks, such as the probability of not achieving the life requirement, were minimized by expensive testing, and those such as launch vehicle failure were minimized by increasing the number of satellites in the program. The important concept was risk reduction, and it was often forced beyond the limits of economic justification due to the psychological effects of a failure. In any case, because of these pressures, the program variables were set easily and it then was the job of the reliability engineer to see that the most reliable satellite was built for the money. Enhancing satellite inherent reliability was often the point of departure for studies in the areas of the parts, test, and redundancy allocation.

In the era of the shuttle the problem becomes more complex. The first question faced is how to measure performance of a satellite program with the addition of resupply. Secondly, what is the proper blend of schedule delay time, resupply frequency, and inherent life which results in the most economic program? Thirdly, how many satellites should be planned for the given mission and what should the selection criteria be? Lastly, the problem also faced without resupply is that of finding the most cost-effective allocation of redundancy which, within the given design, has a specific probability of meeting the life requirements. The models described here were utilized to answer these questions in a quantitative fashion for the Large Space Telescope (LST) satellite program, but they can be applied readily to a general class of programs in the shuttle era. The sequence in which these answers are generated is shown in Fig. 2.

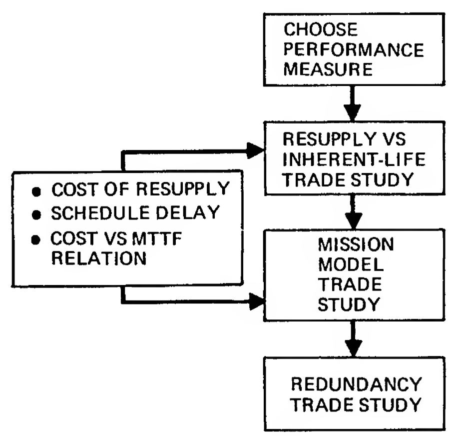


Fig. 2 Elements Required for a Study of a Resupplied Satellite Program

Satellite Program Performance Measure

A variety of reliability measures can be used to determine the performance of a system when resupply is allowed. Among them are system:

- Availability
- Uptime
- Downtime
- Uptime ratio

The definitions of these measures are available in reliability texts[1,2] and are not discussed here. Each measure emphasizes a different facet of the performance of a system, and the choice of a measure should usually be made in terms of system considerations, not merely reliability. In this case, the system was a national space observatory (Fig. 3) which will be used by astronomers to view astronomical phenomena from the highly desirable orbital vantage point. Thus, the astronomers will be paying for high-quality observation time. Therefore, the amount provided them is a measure of observatory performance. In actual practice, the amount of observation time is not exactly equivalent to expected system uptime because of occultation, acquisition, and data transmission, as well as other losses. It is true that maximizing uptime should maximize observation time, and for this reason performance was measured by system uptime.

With the choice of a performance measure completed, strategies for economic optimization had to be developed. From the astronomer's point of view, the strategy was simply to provide the most uptime per dollar (i.e., to minimize the cost per unit of expected satellite uptime). A flaw in this strategy is that it did not take into account funding limitations. Therefore, an additional strategy was employed which determined the lowest cost program achieving a given uptime goal. Using these two approaches, the data generated by the models were evaluated, attractive design regions were identified, and the flexibility gained by the shuttle was easily shown. The shuttle expanded the design region to the point where many alternatives were possible within a given cost range. The best alternatives indicated by both above strategies were within a very close range of design variables.

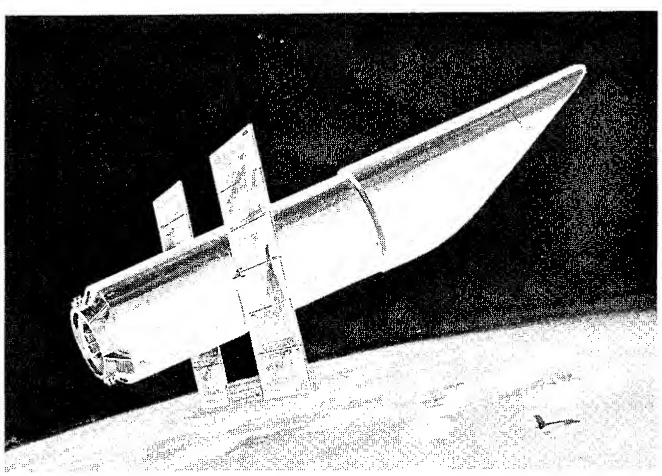


Fig. 3 LST Design Concept Allowing Manned Maintenance

Resupply vs Inherent-Life Trade Study

The next question to be answered subsequent to the choice of a performance measure requires determining the proper blend of resupply and "designed-in" life. The cost of a resupply was determined by the flight and turnaround cost of the shuttle and by refurbishment equipment cost.[3] The cost required to achieve a given life through design is more difficult to estimate. The development of this cost required gathering acquisition cost and life (MTTF) data for similar hardware on several different programs, and is described in detail.[4] The resulting curve is shown in Fig. 4. A rough estimate of the economic decision point for resupply can be obtained by plotting the cost to double the life versus the life itself. This is easily derived from Fig. 4. If we know the approximate cost required for resupply (in our case, $5 million was used), then the life at which this cost occurs becomes the decision point. This trade is shown in Fig. 5. It can be seen from the figure that for any but the smallest life requirements (approximately three months) some benefit is gained through resupply. The cost versus MTTF curve can be used for rough estimating purposes because in many of the cases considered the schedule delay was small compared to satellite life. In this case, a resupply is almost equivalent to adding another satellite in standby, thus doubling the life of the system. Hence, the associated cost can be traded against the cost required to double the system life through design. After the actual model was developed, it was found that the life required to meet the uptime with the lowest cost was approximately one year. As can be seen from Fig. 5, this requirement implies that resupply provides economic benefits, even up to $20 million per flight.

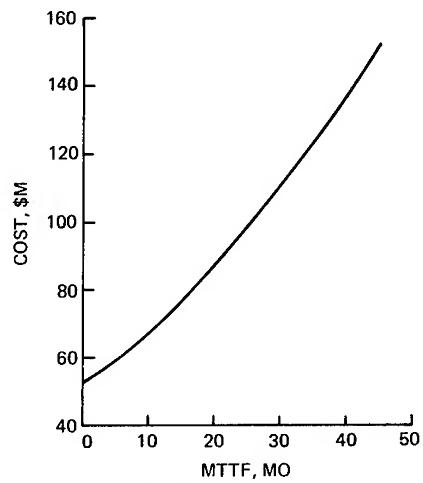


Fig. 4 Satellite Acquisition Cost vs Life

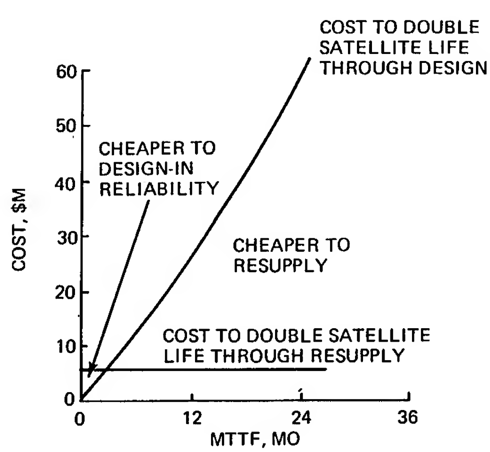


Fig. 5 Resupply/Design Tradeoff

Model for Resupplied System

As mentioned previously, with the addition of resupply many additional programmatic variables must now be included in the optimization of the satellite system. The variables considered for this model are:

- Number of satellites
- Satellite MTTF*
- Shuttle schedule delay (D)
- In-orbit repair proportion (π)
- Survival subsystem MTTF*
- Refurbishment time of satellite on the ground
- Repair time in-orbit (MTTR)

To limit the range of the above variables and the possible system states, the following assumptions were made:

- Only one of every thousand failures which occur cannot be repaired in orbit; that is, only 0.1% of the satellite failure rate would be assigned to non-repairable failures (π = 0.999). (This number was justified by the fact that a concerted effort would be made in the LST design to ensure that all possible failures which occur could be repaired in orbit.)
- A satellite which experiences a non-repairable failure can be retrieved for refurbishment
- The maximum ground refurbishment time required would be one year
- Once a repairable failure occurs, the satellite reverts automatically into a survival mode and awaits refurbishment by the shuttle
- While awaiting a refurbishment flight, the survival subsystem can fail and place the satellite into a catastrophic mode, from which refurbishment is no longer possible
- The probability of failure of the survival subsystem when it is not in operation is zero (pure standby)
- After a shuttle delay period, resupply is initiated if the satellite has not failed catastrophically; the satellite is then returned to the full-up mode

The flow from failure to repair for the one-, two-, and three-satellite cases is shown in Fig. 6, 7, and 8, respectively. The possible conditions in each case are indicated by the system states; each arrow indicates the rate at which a transition from one state to another can occur. For this reason, these illustrations are called state diagrams†. Combinations of system states constitute modes of system operation or mission modes. From the state diagram,

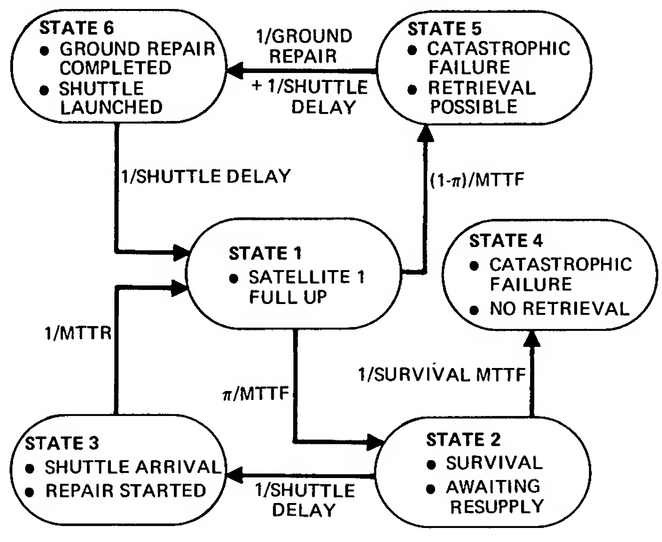


Fig. 6 Single-Satellite State Diagram

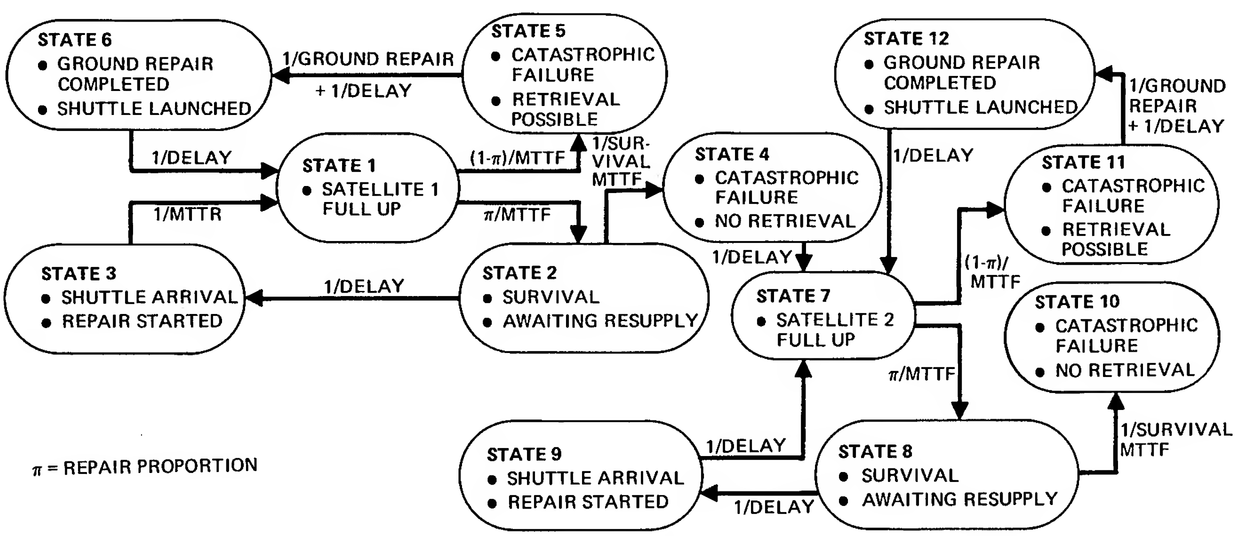


Fig. 7 Two-Satellite State Diagram

---

*Here assumed to be 1/λ by the exponential assumption

†For simplicity the "self-loops," i.e., transitions which result in the same state, were omitted

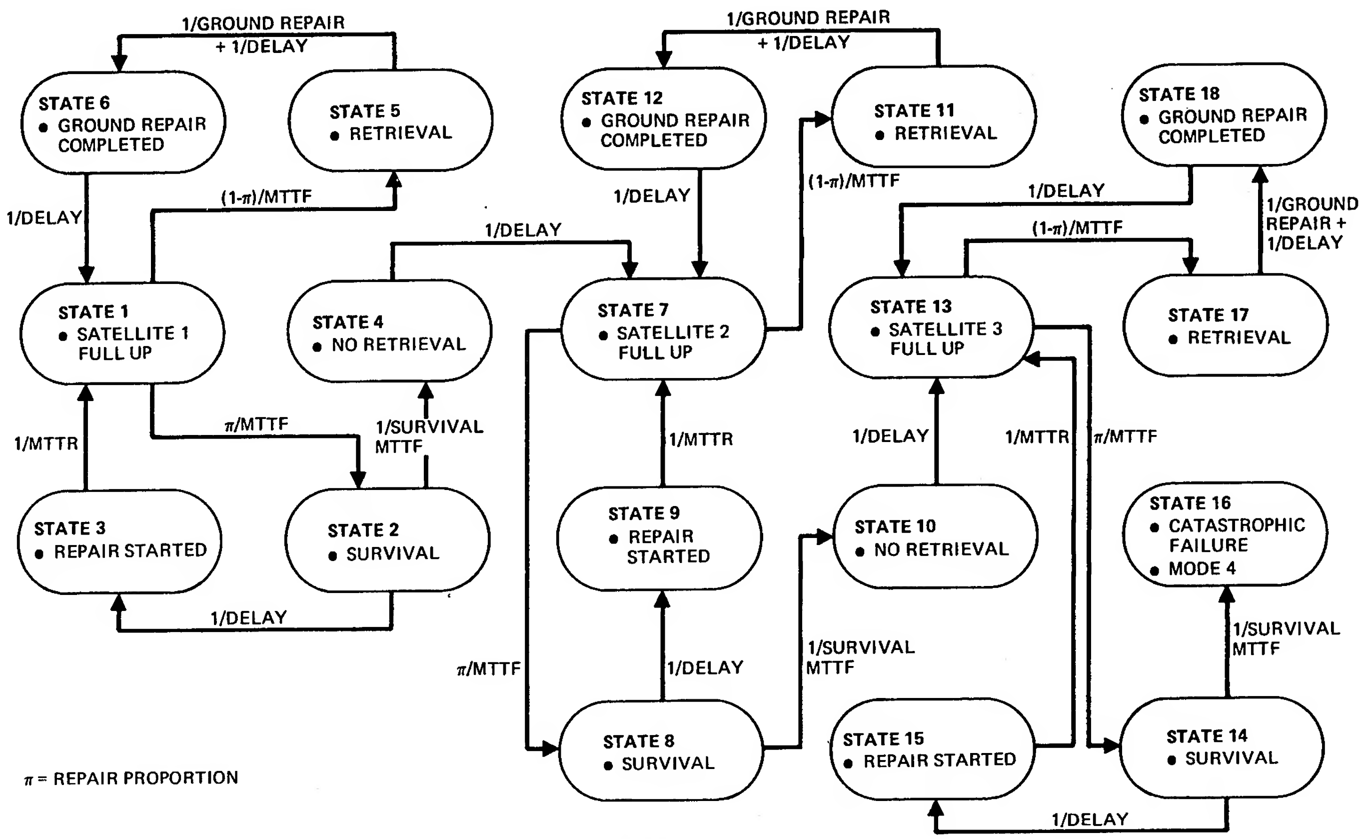


Fig. 8 Three-Satellite State Diagram

a discrete-state, continuous-time Markov model can be constructed to simulate the interactions of the system throughout the mission. Because Markov models are discussed in many references,[1,2,5] our discussion is limited to the exercise of the model.

The flow of the system through time is tracked by changes in the values of the components of a system state vector. Each component of this vector represents the probability that, at a given point in time, the system will be in each particular state.

At the start of the process, the system is assumed to be completely up; thus, the value of the first component of the state vector is one while all other component probabilities are zero. The exact solution for the probability of being in each state at any subsequent time (i.e., the new state vector) is determined by the solution of "n" first-order differential equations, where n is the number of components of the state vector. This problem is extremely difficult, especially when the number of states involved is as large as considered here. Therefore, an approximation method is used, which converges very rapidly to the exact answer. Basically, this solution requires the multiplication of the initial state vector (i.e., the probability of being in each state at time $t = 0$) by the matrix formed by the transition probabilities from one state to another (i.e., the probability of making a transition from one state to another in some small time, Δt). This process produces a new current state vector at time $t + \Delta t$. The multiplication of this current state vector by the transition probability matrix produces a new current state vector at $t + 2\Delta t$. This series of multiplications is continued until the sum of the time increments equals the mission time. The components of the final state vector represents the probability of being in each state at the end of the mission.

The technique just described was developed and programmed using the "Grippo Algorithm"[6] so that a range of transition rates could be handled for each case considered. A sample output of the program called MARKAP is shown in Table 1. The table contains outputs for one-, two-, and three-satellite missions.

Mission Model Trade Study

The output data from Table 1 are plotted in Fig. 9 versus the total program cost (which included approximately $320 million in fixed costs) and uptime. The development of this illustration is explained in detail in References 3 and 4. However, it is obvious that unless expected uptime corrected by the risk incurred is used as a criterion, the best program, in terms of cost and uptime, is that which has only one satellite. The additional satellites achieve these benefits:

- A decrease in the probability of mission termination due to a catastrophic failure
- Greater flexibility in the repair time of a returned vehicle
- Greater flexibility in shuttle response time

The analysis showed that ground refurbishment time was not a driving factor unless refurbishment times greater than satellite life are expected. Thus, the choice of the number of satellites is reduced to a trade between shuttle response delay and the life of the satellite survival subsystem. Some sample results from the model are shown in Fig. 10 and 11. The total results are given in Reference 7. These illustrations show that for a reasonable life survival subsystem (1 year) and a reasonable shuttle delay (0.5 month), the addition of the third satellite does not warrant the increased cost.

Table 1 System Performance Characteristics

Satellite Life, years	Survival S/S Life, years	Shuttle Delay, months	Uptime hours	Uptime Ratio	Probability Up	Probability Catastrophic Failure
ONE-SATELLITE PROGRAM						
0.5	1	0.5	73,400	0.559	0.305	0.667
		1.5	36,963	0.281	0.052	0.935
1	1	0.5	95,958	0.730	0.540	0.437
		1.5	61,499	0.468	0.200	0.777
2	1	0.5	111,580	0.849	0.730	0.254
		1.5	86,374	0.657	0.431	0.544
TWO-SATELLITE PROGRAM						
0.5	1	0.5	113,998	0.868	0.768	0.184
		1.5	82,832	0.630	0.372	0.577
1	1	0.5	123,637	0.941	0.905	0.063
		1.5	106,558	0.811	0.665	0.282
2	1	0.5	127,925	0.973	0.962	0.018
		1.5	120,016	0.913	0.854	0.105
THREE-SATELLITE PROGRAM						
0.5	1	0.5	121,242	0.992	0.908	0.035
		1.5	102,589	0.781	0.654	0.258
1	1	0.5	126,172	0.960	0.959	0.006
			117,154	0.892	0.860	0.070
2	1	0.5	128,681	0.979	0.979	0.001
		1.5	124,102	0.944	0.940	0.013

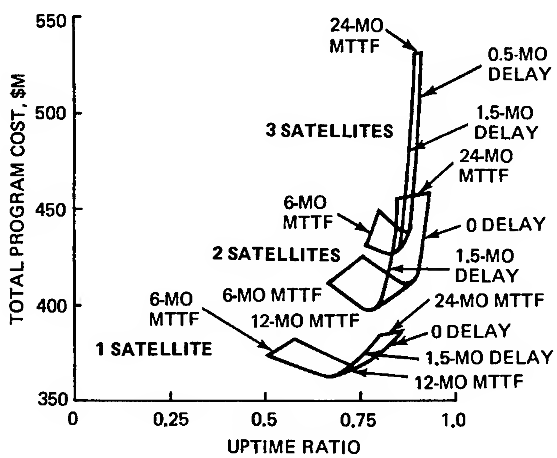


Fig. 9 Cost vs Uptime

The analysis resulted in the following conclusions:

- The two-satellite program was considered most attractive when cost, uptime, and catastrophic failure are considered
- The single-satellite program, although attractive in terms of total program cost, is highly risky in terms of program catastrophic-failure potential
- The two-satellite case is less risky than a single satellite and should be considered optimum in most realistic ranges of program parameters because it provides about the same cost per unit observation time as the single-satellite program
- The three-satellite case does not add enough uptime to make the additional expense attractive; furthermore, this program results in the highest total program cost, as well as the highest cost per unit uptime
- The three-satellite case is relatively insensitive to increases in delay of the shuttle's response and, thus, should be considered as a viable alternative if large delays are expected (e.g., if few shuttles are available)

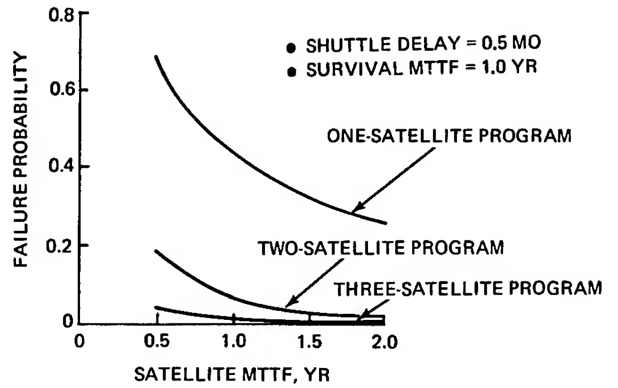


Fig. 10 Effect of Number of Satellites on Program Failure Probability

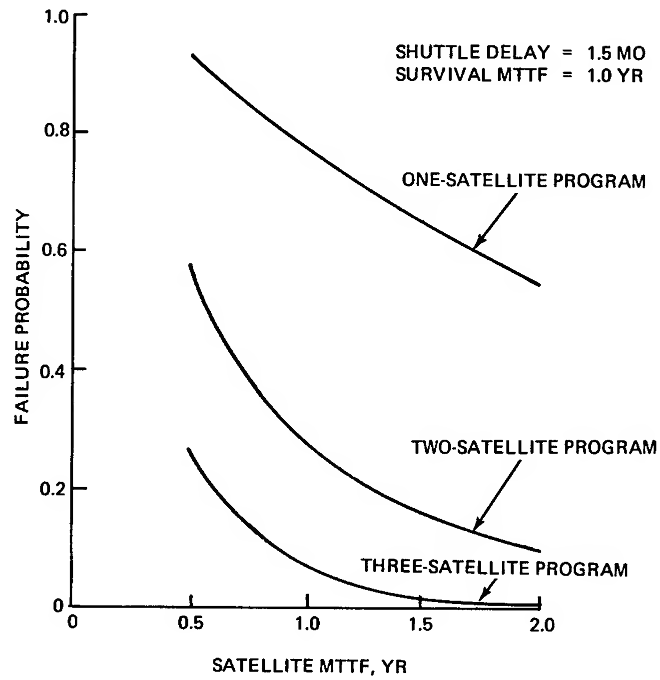


Fig. 11 Effect of Number of Satellites on Program Failure Probability

Due to its favorable expected uptime for most reasonable ranges of program variables, the two-satellite program was recommended for the mission model. The detailed results are shown in Fig. 12. These curves were used to determine the design regions. A simplified version of this set of curves is shown in Fig. 13. The use of these curves to perform tradeoffs is described completely in References

3 and 8. In addition, a brief description is given here. For a given observational time goal, we can determine the lowest-cost program for a given delay by finding the low point on the delay curve. The value of satellite MTTF which produces this minimum should be chosen as the satellite design-life goal. An interesting result which can also be derived from these curves is that the number of failures which would be experienced over a given time can be related to satellite MTTF. Therefore, because each failure requires a shuttle flight, the number of additional flights required for LST repair over its mission life can be determined.

As the simplified curve indicates, the design region lies near a one-year life for high program uptime requirements for both programmatic strategies. Thus, the life requirement for an individual satellite in the program was set at one year.

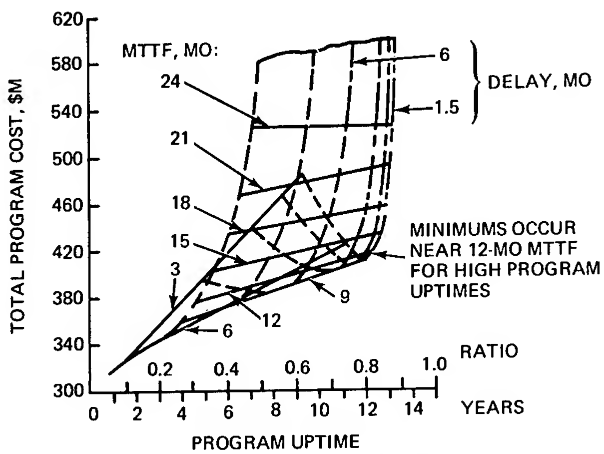


Fig. 12 Cost vs Uptime for Two-Satellite Program

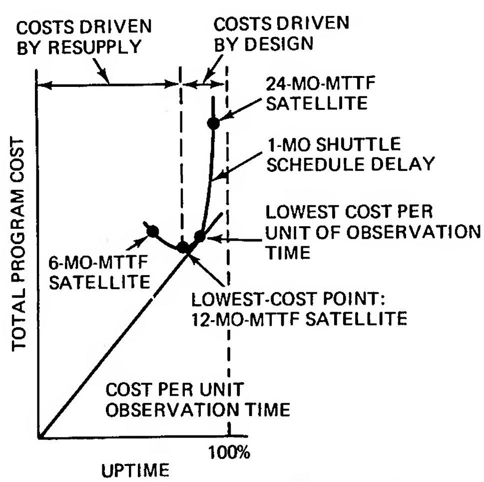


Fig. 13 Optimized Cost of Observation Time

Redundancy Trade Study

With the satellite's life requirement set, the problem becomes, as with a conventional program, one of allocating redundancy in the most effective way. Because it is unreasonable to expect that the baseline configuration will meet the satellite's life requirement, backup systems and/or hardware are required. The question is: Where and how many?

One approach to this is a rule-of-thumb, such as making everything fail-operational and then adding any further redundancy by "guesstimate." Another approach employs the simple method of allocating a portion of the overall requirement approach successively to each system, subsystem, and black box.

While these types of approaches may give reasonable answers, they are generally not optimum. A better solution can be obtained by using the technique of dynamic programming,[9] which yields a truly optimum answer. (In a study of the redundancy allocation on the Orbiting Astronomical Observatory, a possible cost reduction of 33% was found when the results were compared to the initial redundancy allocations.[10])

Dynamic programming is an iterative technique which tests all feasible allocations and selects the optimum. The idea behind such a procedure is very simple. A certain amount of the available resource, cost, is used to add redundant units for a black box. Then the remaining cost is available for allocation to the other boxes. If the number of redundant units added optimizes reliability/cost for that box, then, no matter how the rest of the cost is allocated, this results in the optimum number of units for that box with the specified cost allocated to it. An optimal policy has the property that whatever the initial state and decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.[11] This is the premise behind dynamic programming, and is usually called the "principle of optimality."[12]

The analysis for an entire satellite proceeds as follows. After the system cost limit is chosen, a small cost increment, which is less than the least costly item, must be selected so that no combinations are overlooked. This cost increment is then applied in multiples until the cost limit is reached. The same incremental method is applied to each item in turn. The cost is incremented until a redundant unit of the first item can be added. The resultant increase in reliability is then computed for this addition. The cost continues to be incremented for the first item until the cost constraint is reached. The redundancy allocation which results is the optimum policy for the first item. The same procedure is repeated for the second item. The resulting policy is then compared at each cost increment to the optimum policy for item one and the allocation made to the one which produces the greatest gain in reliability. The resulting allocation is the optimum policy for items one and two combined. This policy is saved and the two single-item allocations discarded. This process is continued using the combined optimum of the previous steps as the basis for obtaining the new optimum policy. After all the items have been considered, the resulting policy is the optimum system reliability within the cost constraint. Because of the way in which the policy is obtained, optimums are available for all system costs from zero to the specified maximum at each cost increment.

The level at which the redundancy is to be added must also be considered. Backup systems (e.g., backup controls) may be incorporated, but, generally, these have different characteristics from the primary system and are added as a matter of policy and not for economic reasons. Redundancy within the units may also be added, but this is difficult and costly once a unit is designed. Thus, the most reasonable level at which to add redundancy to a satellite design is the black box. Because data on these units is usually readily available, the effect of redundancy on the system can be evaluated.

Computerized Optimization

All redundancy is considered to be in standby mode, requiring switching circuitry to be considered. It is expected that switching complexity will limit the increase in reliability due to an added unit to 90% of the incremental gain (based on perfect switching). This factor was included in the expression for standby redundancy. Thus, the reliability of a box with N standby units and failure rate λ is:

$$R(t) = e^{-\lambda t} \sum_{i=0}^{N} \frac{(0.9\ \lambda t)^i}{i!}$$

Recent work[13] proves this formulation, and has been employed in the analysis of fault-tolerant computers using standby systems with imperfect switching.[14]

The analytical technique necessary to perform a dynamic programming optimization has been written for computer solution based upon Reference 15. A simplified flow diagram is presented in Fig. 14. The operational computer program is written in FORTRAN IV and is being used on the IBM 360/75 and 360/67.

The computer first reads the number of items, N, and the total operating time, T. Then, for each item, it reads item cost, failure rate, mode of redundancy, quantity in common, duty cycle, and name. Finally, it reads the redundancy cost constraint and increment. Increment D must be less than or equal to the cost of the lowest-priced item to achieve the exact optimum. The baseline reliability is calculated as:

$$R(t) = e^{-\left(\sum_{i=1}^{N} n_i \lambda_i T_i\right)}$$

where n_i is the number of the i^{th} item required, N is the number of items, and T_i is the operating time of each item in consideration of different duty cycles. The number of cost iterations, Z, is computed as Z = constraint/increment (for integral Z).

The optimization then begins for the first item, i. For each cost iteration, j, ($j \leq Z$), the number of additional units, "X" of item i which can be included at jD cost is found and the standby reliability gain factor,* S_j, is computed, and stored, as:

$$S_j = \sum_{K=0}^{X} \frac{(n_i \lambda_i \delta\ T_i)^K}{K!}$$

where δ is one minus the switching degradation factor, which was taken to be 10%.

------

*Active parallel gain is also calculated using the standard combinational analysis for each redundancy addition to an active parallel unit; but, here, all units were considered to be in standby.

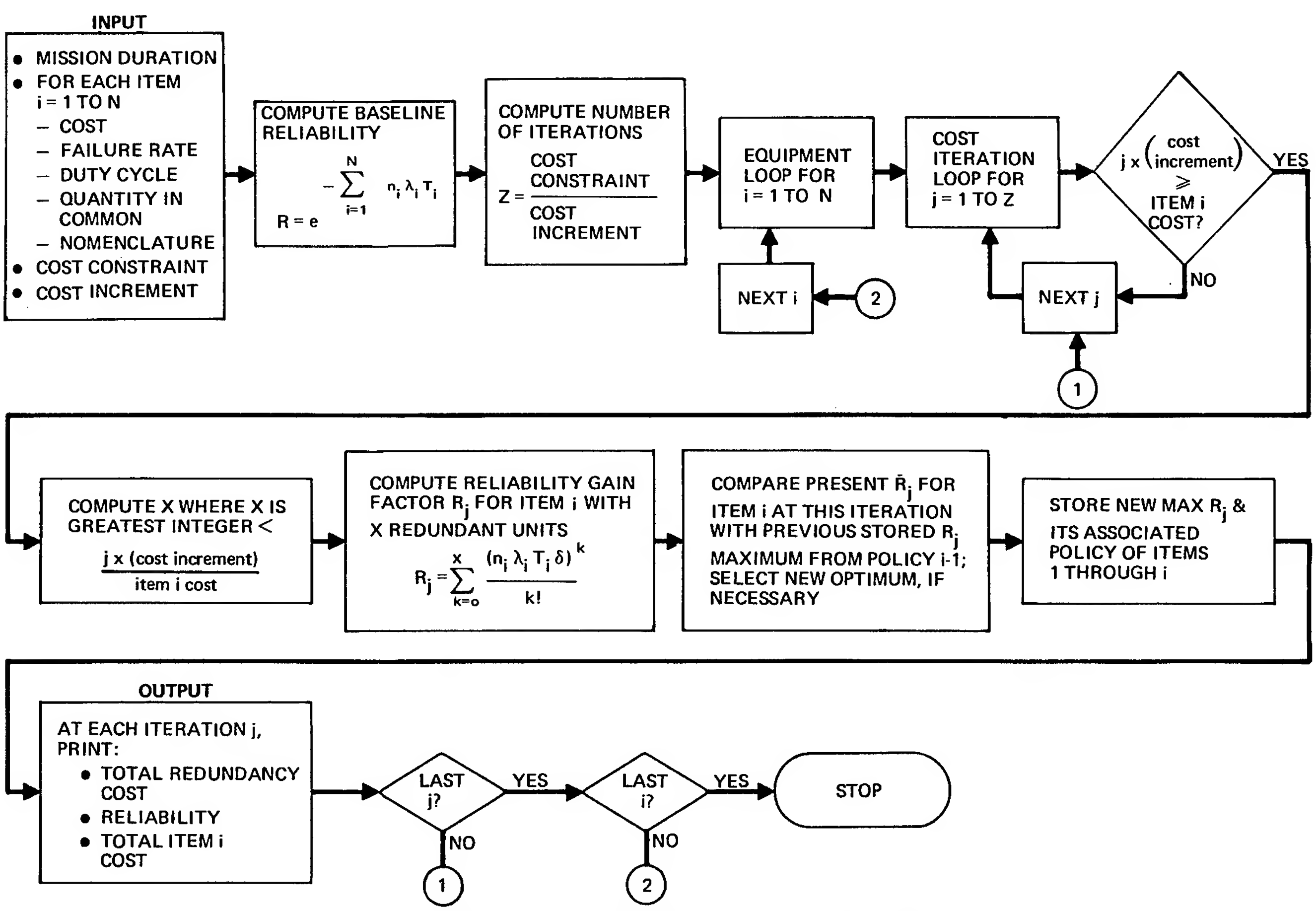


Fig. 14 Logic Flow for Dynamic Programming Redundancy Optimization Technique

In addition, at each iteration J the allocations from the preceding iterations K ($1 \leq K \leq J$) for this item are compared with those from the last stored optimum policy (except during the first-item optimization). All units added in this and the previous policy are then rechecked to determine if a new combination gives a greater gain. This is when some previously added units of the present item may be deleted in favor of a new combination. After Z iterations, the new optimum combination remains stored until a succeeding item optimization requires a new allocation. The final optimum policy is the optimal combination of all items in the configuration.

During each optimization, redundancy cost, reliability, and allotted cost of the present item are printed out at each iteration. The resulting output is a set of N optimization tables, each with Z entries. A sample of the resulting computer run is shown in Fig. 15.

Computer Output Interpretation

To read the optimum redundancy allocation one consults the last optimization table and locates the line where the total cost equals the constraint. On this line is the final achieved reliability and the cost allotted to the last item. The number of unit additions of the last item is obtained from this table. Then the item-allotted cost is subtracted from the constraint and the resource remaining is applied to the preceding optimization table. This remaining redundancy cost is the cost to be allotted to the $(N-1)^{th}$ item. The number of unit additions is found as before. The value in the reliability column of this table is now not required. This process is repeated until the first table is completed and the total redundancy allocation is found.

This process produces a family of optimal allocation policies by starting from any total redundancy cost figure or achieved reliability in the last table and proceeding backward from there. Another feature of this technique is that one may discard the last optimization table and still have the optimum policy for items 1, 2, ..., N-1. In general, for any item i in the sequence, the policy for items 1, 2, ..., i is itself optimum because of the principle of optimality.

Forty-one distinct units on the LST were considered as candidates for additional redundancy. These are shown in Fig. 16. The results are given in detail in Reference 16. A plot of satellite reliability versus cost for a one-year life is shown in Fig. 17. This plot was generated from the data given in the last optimization table alone because, as was explained previously, the final optimization represents the system optimum. It can be seen from this illustration that the marginal cost of an additional increment of reliability increases dramatically between 0.7 and 0.85. Thus, the satellite reliability requirement for a one-year life should be set somewhere between 0.7 and 0.8 to be most economical. At 0.8 reliability, the redundancy policy and additional cost which resulted are given in Table 2.

Conclusions and Extensions of the Analysis

As is true with many economic models, those presented here are dynamic in the sense that the preliminary design variables obtained are fed back into the models with further design refinements to produce more definitive design requirements. In this case, the tools presented were successful in determining the number of satellites to be used in the mission model, and the region in which design life and reliability requirements could be found for an individual satellite in the program. However, the design presented for redundancy optimization, although reasonable, is probably significantly different from the final functional design. In some cases, the redundancy indicated by dynamic programming is not realistic in that the configuration of redundancy recommended is not realizable from a functional viewpoint. In spite of these drawbacks, even this "first-cut" analysis is extremely useful because it highlights for the design engineer where the weak links are in his preliminary design. As such, the analysis directs his attention to improving those areas by hardware replacement or redesign, allowing the design to mature in the most cost-effective manner.

OPTIMIZATION TABLE... 41
ITEM COST... 65448.10 HARN.PYRO.HTRS.ANT EPS

TOTAL COST	RELIABILITY	ALLOCATION	REMAINING	NO.	FEC. COST
0.0	0.17042506	0.0	0.0	0	0.0
10000.0	0.17311215	0.0	10000.0	0	0.0
20000.0	0.18472117	0.0	20000.0	0	0.0
30000.0	0.18763351	0.0	30000.0	0	0.0
40000.0	0.20382768	0.0	40000.0	0	0.0
50000.0	0.20704138	0.0	50000.0	0	0.0
60000.0	0.22092569	0.0	60000.0	0	0.0
70000.0	0.22440904	0.0	70000.0	0	0.0
80000.0			80000.0	0	0.0
[illegible]	0.83777094	[illegible]	[illegible]	0	0.0
[illegible]	0.83777094	70000.0		0	0.0
9860000.0	0.83777237	70000.0	[illegible]		0.0
9870000.0	0.83777237	70000.0	9800000.0	[illegible]	[illegible]
9880000.0	0.83777303	70000.0	9810000.0	1	
9890000.0	0.83777368	70000.0	9820000.0	1	[illegible]
9900000.0	0.83777368	70000.0	9830000.0	1	65448.1
9910000.0	0.83777452	70000.0	9840000.0	1	65448.1
9920000.0	0.83777452	70000.0	9850000.0	1	65448.1
9930000.0	0.83777452	70000.0	9860000.0	1	65448.1
9940000.0	0.83777452	70000.0	9870000.0	1	65448.1
9950000.0	0.83777452	70000.0	9880000.0	1	65448.1
9960000.0	0.83777452	70000.0	9890000.0	1	65448.1
9970000.0	0.83777452	70000.0	9900000.0	1	65448.1
9980000.0	0.83777452	70000.0	9910000.0	1	65448.1
9990000.0	0.83777452	70000.0	9920000.0	1	65448.1
10000000.0	0.83777469	70000.0	9930000.0	1	65448.1

Fig. 15 Dynamic Programming Output Sample

ITEM NO.	COST	LAMBDA	CTY	D/CYCLE	CODE	NAME	
1	68304.4	0.1500E-04	1	1.00	0	FIXED HEAD TRACKER	S&C
2	9935.2	0.6900E-07	1	1.00	0	DIGITAL SUN SENSOR	S&C
3	31047.5	0.2486E-04	1	1.00	0	DIGITAL SUN SENSOR ELECT	S&C
4	931424.9	0.5500E-05	1	1.00	0	I R U + ELECTRONICS	S&C
5	[illegible]	0.6500E-06	3	[illegible]	0	FINE WHL & JET CONTR	S&C
[illegible]	397405.0	0.5870E-06	1	1.00	0	MAGNETOMETER	S&C
38	8382.8	0.2000E-05	1	1.00	0	MULTIPLEXER	S&C
39	37257.0	0.7900E-05	1	1.00	0	DIODE BOX	EPS
40	18628.5	0.2500E-05	1	1.00	0	POWER DIST UNIT	EPS
41	65448.1	0.1000E-06	1	1.00	0	HARN,PYRO,HTRS,ANT	EPS

THE COST LIMIT IS 0.1000E 08
THE INCREMENT IS 10000.
THE OPERATING TIME IS 8760.0

Fig. 16 LST Dynamic Programming Input Data

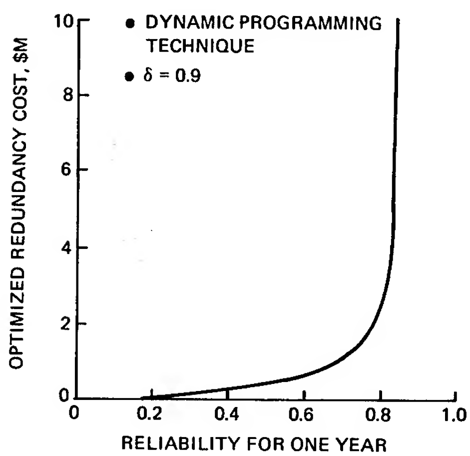


Fig. 17 LST Cost of Redundancy to Achieve Reliability

Design maturity is also greatly aided by a further application of the MARKAP program. Once the mission model is selected, the scant number of states can be expanded in the chosen state diagram to include many degraded modes of satellite operation as well as the recommended redundancy. The degraded modes are defined by grouping various hardware within the design under headings describing the operational performance of the satellite with these units failed or degraded. Then, additional system states are added to the Markov model so that sequence of hardware failures within a group is adequately represented. The sum of the probabilities in each of the states composing a mode at the end of the mission time is the mission mode probability. By assigning a value to the quality of the observational capability of the satellite in each of the degraded modes, the most effective observing system can be obtained. This is accomplished by determining the mixture of hardware failure rates within each group which produces the best system within the total program cost constraint. The second iteration is reflected in design changes, which can be again analyzed using dynamic programming to produce even more specific design recommendations.

Thus, the tools presented here are truly dynamic in that they can be applied at successive stages in the design process to produce pertinent design recommendations for each stage.

Table 2 LST Redundancy Recommendation for 1-Year Design Life, 0.8 Reliability[1]

Subsystem	Item	Qty[2]	Cost[3], $K
Stabilization & Control	Fixed Head Tracker	1	68.30
	Digital Sun Sensor Electronics	2	62.10
	IRU & Electronics	1	931.42
	Fine Wheel & Jet Controller	3	620.95
	Magnetometer	1	18.63
	Wheels	1	124.19
	Remote Decoder	1	19.56
	Multiplexer	1	8.38
	Wiring Harness	1	14.22
	Magnetometer Electronics	1	15.52
	SAS Electronics	2	37.26
Communications & Data Handling	Command Receiver	1	136.61
	Narrowband Transmitter	1	37.26
	Command Decod-Detect-Verif	1	19.56
	Telem Format Controller	1	55.89
	Computer Ops Monitor	1	24.84
	Wideband Transmitter	1	93.14
	Power Amp & RF Switch	1	12.42
	Multiplexer	1	8.83
	Power Converter	1	18.63
Pneumatics	Gas Tanks	1	7.45
	Jets, Connectors, & Solenoid Valve	1 Set	8.56
Electrical Power	Multiplexer	1	8.38
	Diode Box	1	37.30
	Power Distribution Unit	1	18.60
TOTAL REDUNDANCY COST			$2.41 M

1. Only hardware requiring unit redundancy are presented.
2. In addition to baseline.
3. Including G&A and fee.

Acknowledgement

The authors are grateful to Mr. R. Farkas for his large contribution in the performance of the analysis on which this paper is based, as well as to Messers M. Olstad, F. Kratochvil, and W. Carroll for their suggestions which greatly improved the quality of this paper.

References

1. Shooman, M. L., Probabilistic Reliability: An Engineering Approach, McGraw-Hill, Inc., New York (1968)

2. Rau, J. G., Optimization and Probability in Systems Engineering, Van Nostrand Reinhold Company, New York (1970)

3. "OAO/LST Shuttle Economics Study," Grumman Aerospace Corporation, Contract NAS 5-17149, Goddard Space Flight Center, October 1970

4. Diamond, E. H., and Fragola, J. R., "Satellite Long-Life Assurance: The Impact of the Shuttle Era," AIAA, Man's Role in Space Conference, Cocoa Beach, Florida, March 27-28, 1972

5. Messinger, M., "Models Analysis, and Approximations for System Reliability and Availability," AFSOR 67-1438, Contract AF-49(638)-1402, June 1967

6. Grippo, G., "Evaluation of Reliability and Availability Measures for Complex Systems," IEEE Reliability Symposium, Reading, Mass., April 1969

7. Fragola, J. R., and Farkas, R., "Large Space Telescope Mission Model Optimization," Grumman Memo LST-ENG-71-051, June 30, 1971

8. Salee, A., "An Analytic Technique to Assess the Economic Impact of the Shuttle on Satellite Payloads," AIAA Space Systems Meeting, Denver, Colorado, July 19-20, 1971

9. Bellman, R., and Dreyfus, S., Applied Dynamic Programming, Princeton University Press, Princeton, New Jersey (1960)

10. Fragola, J. R., and Farkas, R. J., "OAO/Shuttle Economics Study Redundancy - Cost Optimization," Grumman Memo LST-ENG-71-056, July 8, 1971

11. Messinger, M., and Shooman, M. L., "Techniques for Optimum Spares Allocation: A Tutorial Review," IEEE Transactions on Reliability, Vol. R-19, No. 4, Nov. 1970, p. 157

12. Derman, C., Finite State Markovian Decision Processes, Academic Press, New York (1970)

13. Bouricius, W. G., et al, "Reliability Modeling Techniques for Self-Repairing Computer Systems," Proc. ACM 1969 Annual Conference

14. Bouricius, W. G., et al, "Reliability Modeling for Fault-Tolerant Computers," IEEE Transactions on Computers, Vol. C-20, No. 11, November 1971

15. Shooman, M., "Spares Provisioning and Reliability Optimization for the Space Station Project," Unpublished Grumman Report, July 1967

16. Fragola, J. R., and Farkas, R., "Large Space Telescope (LST) Redundancy - Cost Optimization," Grumman Memo LST-ENG-71-058, July 14, 1971

THE KOLMOGOROV-SMIRNOV TEST MODIFIED FOR CENSORED DATA

INDEX SERIAL NUMBER — 1063

Gerald E. Sasser, Jr.
General Manager, McDonald's Restaurant
Provo, Utah 84601

Dale O. Richards
Professor of Statistics
Brigham Young University
Provo, Utah 84601

Abstract

Due to the occurrence of censored data in life testing, it is necessary to modify the original Kolmogorov goodness-of-fit test to provide the correct significance level. In this paper, two alternative procedures are offered: a symmetric test and a central, non-symmetric test. The theory underlying the modifications is provided along with an example of the use of the tests.

I. Censored Data in Life Testing

One problem which arises often in the field of life testing is a sample which is incomplete due to censoring or truncation. A sample is said to be censored when the experiment is terminated after a given number of observations. It is said to be truncated when the experiment is terminated at a given point in time. If the researcher is faced with such a sample he must be aware of what happens to the significance level and the power of the test when the full sample tables are used to determine a critical value for λ. This paper presents a solution to the problems caused by censored or truncated samples.

II. The Modified Test

The standard Kolmogorov-Smirnov test does not take into consideration the fact that, in general, negative values of $S_N(x)-F(x)$ are a little more likely due to the nature of the empirical distribution function. For this purpose, we present a test which uses different positive and negative critical values allowing α to be evenly divided on the two sides, as well as a typical symmetric test. We also note that positive differences seem more likely at the lower end of the distribution while negative differences should occur more frequently at the upper end. We are currently working on a procedure which will include the necessary adjustments to handle this situation. It is hoped that this new procedure will enhance the power of the test by a significant amount.

A. The Symmetric Test

For each i = 1, 2, ..., N-1, there is a value x_i such that $F(x_i)=i/N$. (If $F(x)=i/N$ on some interval, let x_i be the left endpoint of that interval.) Let x' be the value after which the data are unavailable. Let x_k be the particular x_i such that $x_{k-1}<x'\leq x_k$. Define D_k to be sup $|S_N(x)-F(x)|$ for all $x\leq x_0'$. Then critical values for D_k for any desired significance level can be obtained by means of Theorem 3 in part III of this paper. In practice, it would perhaps be easier to use the charts in Appendix 1 which were obtained by means of Theorem 3.

B. The Central Non-symmetric Test

Let $D_k^+=\sup |S_N(x)-F(x)|$ when $S_N(x)\geq F(x)$ and let $D_k^-=\sup |S_N(x)-F(x)|$ when $S_N(x)\leq F(x)$. The test procedure is the same as for the symmetric test but different critical values are used for D_k^+ and D_k^-. In this test, the critical values are obtained by setting the appropriate probability formulae equal to $\alpha/2$. (These formulae are found in Theorems 1 and 2 in Part III.) Charts for these values are also provided in Appendix 1.

C. Example

This example comes from a report entitled "Tests for the Validity of the Assumption that the Underlying Distribution of Life is Exponential" by Benjamin Epstein.[1] In the example, the hypothesis to be tested is that the data come from a uniform distribution with $F(x)=x/1363$. For the example, N=50 and $\alpha=.05$. To illustrate the use of the above tests the sample was assumed to have been censored at $x'=600$. This means that k=23 since $.44<F(600)<.46$. The results of the standard Kolmogorov-Smirnov test and the tests for censored data are shown in Figure 1. Note that the significance level for the censored data case is much lower than .05 when the standard tables are used.

III. Theory

The theorem underlying the standard test was first proved by A. N. Kolmogorov.[4] There have been several attempts to simplify this original proof. One of those attempts was presented by William Feller.[2] The theory about to be presented relies heavily on the techniques used by Feller.

A. Theorems

Theorem 1: Let $F(x)$ be a continuous cumulative distribution function and let D_k^+ be defined as in Part II. Then, as $N\to\infty$,

$$\Pr(D_k^+>\lambda N^{-1/2}) \to L^+(\lambda) \text{ where}$$

$$L^+(\lambda)=N\lambda \sum_{r=1}^{k} \left\{ \frac{(N-r)^{N-r-\lambda N^{1/2}}}{\Gamma\,(N-r-\lambda N^{1/2}+1)} \; \frac{e^{r-N}}{r^{3/2}} \sum_{\ell=1}^{\infty} (-1)^{\ell-1} (2\ell-1)e^{\frac{-\lambda^2 N(2\ell-1)^2}{2r}} \right\}$$

The research for this paper was supported in part by USAF Contract No. F42600-72-C-0016.

Theorem 2: Let $F(x)$ be a continuous cumulative distribution function and let D_k^- be defined as in Part II. Then, as $N \to \infty$,

$$Pr(D_k^- > \lambda N^{-1/2}) \to L^-(\lambda) \text{ where}$$

$$L^-(\lambda)=N\lambda \sum_{r=1}^{k} \left\{ \frac{(N-r)^{N-r+\lambda N^{1/2}}}{\Gamma(N-r+\lambda N^{1/2}+1)} \frac{e^{r-N}}{r^{3/2}} \sum_{\ell=1}^{\infty} (-1)^{\ell-1} (2\ell-1)\, e^{\frac{-\lambda^2 N(2\ell-1)^2}{2r}} \right\}$$

Theorem 3: Let $F(x)$ be a continuous cumulative distribution function and let D_k be defined as in Part II. Then, as $N \to \infty$,

$$Pr(D_k > \lambda N^{-1/2}) \to L(\lambda) \text{ where}$$

$$L(\lambda) = L^+(\lambda) + L^-(\lambda)$$

B. Proof

Since $F(x)$ is continuous, it is possible to define numbers x_i such that

(1) $F(x_i)=i/N \quad (i=1,2,\ldots,N-1)$.

This definition is unique except when $F(x)=i/N$ within an entire interval, in which case we define x_i as the left endpoint of that interval.

Let $\lambda N^{-1/2}$ be the sup $|S_N(x)-F(x)|$ for all x in the range of the censored sample. Let c be the greatest integer less than or equal to $\lambda N^{1/2}$.

Then the inequality

(2) $S_N(x)-F(x)>\lambda N^{-1/2}$

implies that

(3) $S_N(x)-F(x)>c/N$.

Inequality (3) will hold within some maximal interval. At the right endpoint, Ψ, of this interval we have

(4) $S_N(x)-F(x)=c/N$

Since $S_N(\Psi)=r/N$ for some integer r and c has been defined as an integer, we have

(5) $F(\Psi)=i/N$ and $\Psi=x_i$ for some i.

Thus, $X^*_{i+c} \le x_i < X^*_{i+c+1}$, or in other words: exactly i+c among the N variables are smaller than x_i. We denote this event by $A_i(c)$. Then (2) holds if and only if at least one of the events $A_1(c)$, $A_2(c)$, ..., $A_k(c)$ occurs. The argument applies equally to c<0 and shows that the event $D_k>\lambda N^{-1/2}$ occurs if and only if at least one of the events

(6) $A_1(c)$, $A_1(-c)$, $A_2(c)$, $A_2(-c)$, ..., $A_k(c)$, $A_k(-c)$ occurs.

Let U_r and V_r be the events that in the sequence (6) the first events to occur are $A_r(c)$ and $A_r(-c)$ respectively. These events are mutually exclusive and therefore,

(7) $Pr(D_k>\lambda N^{-1/2})= \sum_{r=1}^{k} Pr(U_r)+ \sum_{r=1}^{k} Pr(V_r)$

From the definitions of the terms involved we have

(8) $Pr\{A_i(c)\}= \sum_{r-1}^{i} Pr(U_r)Pr\{A_i(c)|A_r(c)\}+ \sum_{r=1}^{i} Pr(V_r)Pr\{A_i(c)|A_r(-c)\}$

and

$Pr\{A_i(-c)\}= \sum_{r=1}^{i} Pr(U_r)Pr\{A_i(-c)|A_r(c)\}+ \sum_{r=1}^{i} Pr(V_r)Pr\{A_i(-c)|A_r(-c)\}$.

This is a system of 2k linear equations for the 2k unknowns $Pr(U_r)$ and $Pr(V_r)$ and we will solve it by the method of generating functions.

By definition of x_i, $Pr(X_m<x_i)=i/N$. Then the probability of the event $A_i(c)$ is given by

(9) $Pr(A_i(c))=\binom{N}{i+c}\left(\frac{i}{N}\right)^{i+c}\left(\frac{N-i}{N}\right)^{N-i-c}$

Similarly, for r<i,

(10) $Pr(A_i(c)|A_r(c))=\binom{N-r-c}{i-r}\left(\frac{i-r}{N-r}\right)^{i-r}\left(\frac{N-i}{N-r}\right)^{N-i-c}$

and

(11) $Pr(A_i(c)|A_r(-c))=\binom{N-r+c}{i-r+2c}\left(\frac{i-r}{N-r}\right)^{i-r+2c}\left(\frac{N-i}{N-r}\right)^{N-i-c}$

The probabilities for $A_i(-c)$ can be found in a similar fashion. (9), (10), and (11) can be written more conveniently in terms of the quantities

(12) $p_i(c)= \frac{e^{-i}\, i^{i+c}}{(i+c)!}$

We then have

(13) $Pr(A_i(c))=p_i(c)p_{N-i}(c)/p_n(0)$

(14) $Pr(A_i(c)|A_r(c))=p_{i-r}(0)p_{N-i}(-c)/p_{N-r}(-c)$

(15) $Pr(A_i(c)|A_r(-c))=p_{i-r}(2c)p_{N-i}(-c)/p_{N-r}(c)$

We may then use these to simplify (8) so that

(16) $p_i(c)/p_N(0)= \sum_{r=1}^{i} Pr(U_r)p_{i-r}(0)/p_{N-r}(-c) + \sum_{r=1}^{i} Pr(V_r)p_{i-r}(2c)/p_{N-r}(c)$

and

$p_i(-c)/p_N(0)= \sum_{r=1}^{i} Pr(U_r)p_{i-r}(-2c)/p_{N-r}(-c) + \sum_{r=1}^{i} Pr(V_r)p_{i-r}(0)/p_{N-r}(c)$

Let

(17) $u_r=Pr(U_r)p_N(0)/p_{N-r}(-c)$ and

$v_r=Pr(V_r)p_N(0)/p_{N-r}(c)$

Then (16) further simplifies to

(18) $p_i(c)= \sum_{r=1}^{i} u_r p_{i-r}(0) + \sum_{r=1}^{i} v_r p_{i-r}(2c)$

and

$p_i(-c)= \sum_{r=1}^{i} u_r p_{i-r}(-2c)+ \sum_{r=1}^{i} v_r p_{i-r}(0)$

This system is of the convolution type and can therefore be solved by means of generating functions. Let

(19) $u(\omega)=\sum_{i=1}^{\infty} u_i\omega^i$ and $v(\omega)=\sum_{i=1}^{\infty} v_i\omega^i$

and

(20) $p(\omega;c)=N^{-1/2}\sum_{i=1}^{\infty} p_i(c)\omega^i$

Then (18) reduces to

(21) $p(\omega;c)=u(\omega)\ p(\omega;0) + v(\omega)\ p(\omega;2c)$

and

$p(\omega;-c)=u(\omega)\ p(\omega;-2c) + v(\omega)\ p(\omega;0)$

We now let $c \to \infty$ and $N \to \infty$ so that $\lambda = (c+\epsilon)N^{-1/2}$ remains constant.
Then

(22) $N^{1/2}p_i(\lambda) \to (2\pi t)^{-1/2}\exp(-\lambda^2/2t)$

as $k/N \to t$. Using (20) and (22) the continuity theorem[1] implies that

(23) $p(e^{-s/N}\ \lambda N^{1/2}) \to (2s)^{-1/2}\exp(-(2s\lambda^2)^{1/2})$

Solving the system (21) for $u(\omega)$ and $v(\omega)$ we find

(24) $\lim_{N\to\infty} u(e^{-s/N})=\lim_{N\to\infty} v(e^{-s/N})=\dfrac{\exp(-(2s\lambda^2)^{1/2})}{1+\exp(-(8s\lambda^2)^{1/2})}$

Let $\alpha_i = Nu_i$.
Then

(25) $\frac{1}{N}\ \alpha(e^{-s/N}) \to \sum_{\ell=1}^{\infty} (-1)^{\ell-1}\ \exp((-(2l-1)\sqrt{2}\lambda)s^{1/2})$

From the continuity theorem,

(26) $f(t)=\frac{1}{\sqrt{2\pi}}\ \frac{\lambda}{t^{3/2}}\sum_{\ell=1}^{\infty}(-1)^{\ell-1}(2\ell-1)\exp$

$(-\lambda^2(2\ell-1)^2/2t)$

and

(27) $\alpha_i \to f(i/N)=\frac{1}{\sqrt{2\pi}}\ \frac{\lambda N^{3/2}}{i^{3/2}}\sum_{\ell=1}^{\infty}(-1)^{\ell-1}(2\ell-1)\exp$

$(-(\lambda N^{1/2})^2(2\ell-1)^2/2i)$

Since $\alpha_i = Nu_i$,

(28) $u_i \to \frac{1}{\sqrt{2\pi}}\ \frac{\lambda N^{1/2}}{i^{3/2}}\sum_{\ell=1}^{\infty}(-1)^{\ell-1}(2\ell-1)\exp$

$(-(\lambda N^{1/2})^2(2\ell-1)^2/2i)$

Then, from (17),

(29) $\Pr(U_r) \to \dfrac{N!(N-r)^{N-r-\lambda N^{1/2}}}{N^N\Gamma(N-r-\lambda N^{1/2}+1)}\ \dfrac{\lambda N^{1/2}e^r}{\sqrt{2\pi}r^{3/2}}\sum_{\ell=1}^{\infty}(-1)^{\ell-1}$

$(2\ell-1)\exp\left(\dfrac{-(\lambda N^{1/2})^2(2\ell-1)^2}{2r}\right)$

[1]Continuity Theorem: If, as $\delta\to 0$, $\delta\ u(e^{-\delta s})\to\emptyset(s)$, then, for every fixed $t>0$, $u\to f(t)$ when $k\ \delta\to t$; conversely, if $u_k\to f(t)$ when $k\ \delta\to t$, then $\delta\ u(e^{-\delta s})\to\emptyset(s)$. $\emptyset(s)$ is the Laplace transform of $f(t)$.

and finally, using Sterling's approximation for N!,

(30) $\Pr(U_r) \to \dfrac{N\lambda(N-r)^{N-r-\lambda N^{1/2}}}{\Gamma(N-r-\lambda N^{1/2}+1)}\ \dfrac{e^{r-N}}{r^{3/2}}\sum_{\ell=1}^{\infty}(-1)^{\ell-1}$

$(2\ell-1)\exp\left(\dfrac{-(\lambda N^{1/2})^2(2\ell-1)^2}{2r}\right)$

The $\Pr(V_r)$ follows thru in the same manner. Hence all three theorems are proved.

Summary of Conclusions

The Kolmogorov goodness-of-fit test, when applied to censored data, provides a critical value too large for the quoted significance level. This, of course, reduces the power of the test unnecessarily. By using the modification offered here it is possible to obtain the correct critical value for the quoted significance level.

Appendix

An understanding of the research described in this paper requires some knowledge of tests of hypothesis and the standard Kolmogorov-Smirnov goodness-of-fit test.

A. Tests of Hypothesis

The need to test the validity of a given hypothesis often arises in the field of inferential statistics. The general procedure consists of the following steps.

First, the hypothesis and any alternatives of interest are formally stated. Then, a rule is formulated to determine whether or not to reject the hypothesis based on the results obtained in taking a random sample of the population in question. The sample is then taken and analyzed in order to make the final decision.

The rule described above is usually based on the probabilities of making the two possible types of error. A Type I error occurs when the hypothesis stated is true but is rejected. A Type II error occurs when the hypothesis is false but is not rejected. The probability of a Type I error is usually referred to as the size or significance level of the test and is usually referred to as the size or significance level of the test and is usually represented by the Greek letter α. The probability of a Type II error is represented by the Greek letter, β, and $1-\beta$ is called the power of the test, i.e., the probability of rejecting a false hypothesis. The rule usually seeks to minimize the probabilities of these errors and the sample size.

B. The Standard Kolmogorov-Smirnov Goodness-of-fit Test

Let $X_1, X_2, \ldots, X_N$ represent a random sample of size N from a population with hypothesized cumulative distribution function, $F(x)$, and let $F(x)$ be a continuous function. Let $X_1^*, X_2^*, \ldots, X_N^*$ be this same sample reordered so that $X_1^* \le X_2^* \le \ldots \le X_N^*$. The empirical distribution function is defined to be:

$$(1)\ S_N(x) = \begin{cases} 0 & \text{if } x < X_1^* \\ k/N & \text{if } X_k^* \le x < X_{k+1}^* \\ 1 & \text{if } X_N^* \le x \end{cases}$$

The statistic used to test the hypothesis is, then,

$$(2)\ D_N = \sup |S_N(x)-F(x)|\ ^2$$

Tables have been formulated showing $\Pr(N^{1/2}D_N>\lambda)$ when the hypothesis is true. The acceptable significance level is used to determine the critical value of λ so that, if $D_N>\lambda N^{-1/2}$, the hypothesis is rejected. The power of this test is notoriously small as is the power of most distribution-free procedures.

Another distribution-free procedure sometimes used is the chi-square test. Many authors agree, however, with E. S. Keeping[3] who states, "The power of the chi-square test in general is not known, but in some cases where comparison with the Kolmogorov test is possible, it appears that the latter is much the more powerful of the two." The advantage of the chi-square test is that the effect of using estimated parameters is known to merely reduce the degrees of freedom. The effect this would have on the Kolmogorov-Smirnov test is not really known.

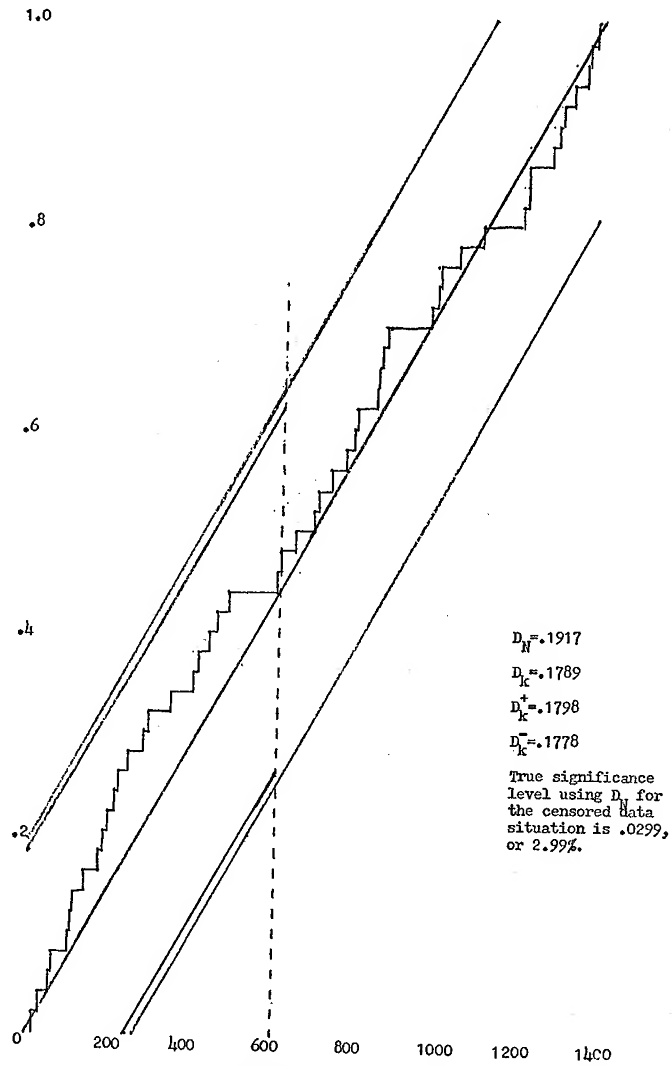


FIGURE 1.

Bibliography

1. Epstein, Benjamin. Statistical Techniques in Life Testing, Wayne State University, Detroit, Michigan, 1959.
2. Feller, William. "On the Kolmogorov-Smirnov Limit Theorems for Empirical Distributions," Annals of Mathematical Statistics, 1948.
3. Keeping, E. S. Introduction to Statistical Inference.
4. Kolmogorov, A. N. "Sulla Determinazione Empirica di Una Legge di Distributione," Inst. Ital. Attuari, Giorn., Vol. 4(1933), pp. 1-11.

[2] The term "sup" is short for the word "supremum" and is used here because a maximum value for $|S_N(x)-F(x)|$ does not always exist. However, a "least upper bound" does always exist and the "sup" is then equal to this bound.

MECHANICAL RELIABILITY I AND II

Session Organizers' Report

Gerhard Reethof
The Pennsylvania State University
University Park, Pa., 16802

In preparing myself for the writing of these introductory remarks to the Mechanical Reliability Sessions of this Conference, I reviewed my remarks in the 1969, 1970, and 1971, Proceedings of the Reliability and Maintainability Conferences. I can say, without reservation, that the technical level of the papers continues to rise, and that we continue to receive a substantial number of papers from the academic world, with many of the contributions representing work with industry. This trend of contributions from academe in our field, gives clear indications that the concepts and techniques of particularly mechanical structural reliability are finding their way into engineering education. This is, of course, an important development since mechanical engineering education continues to be weak in the areas of probabilistic design methods and reliability/maintainability techniques.

As we look back over the last ten years, we see this increasing participation from "the Professors." Back in 1965, for example, Volume 4 of the annuals of Reliability and Maintainability, we find only 2 papers from academe out of a total of 88 papers. Yet, in soliciting the papers for this year's conference, I am finding increasing reluctance on the part of the engineers in industry to report on some of their more recent advances in the state-of-the-art of mechanical reliability. We find, therefore, that the majority of the 2 sessions' papers came from the academicians, my urgent pleas not withstanding. I continue to be fully aware of some of the very excellent work going on within several major corporation on material strength characteristics and mechanical reliability techniques. May I be permitted to urge these investigators to expose their excellent work to the engineering public so that all of us can gain in insight and an improved technology. I should hasten to add in this context, that we are most fortunate in having Trevor Salt's paper presented to us, which I believe will represent quite a remarkable milestone in new and better understanding of the fatigue failure mechanism, cumulative damage modeling and crack inception, as well as propagation in fracture mechanics.

Continuing to discuss the Mechanical/Reliability Program at this Conference, you will note that we have two sessions which is, of course, the result of the submittal of, what I hope you will agree, many excellent papers. The papers again cover a wide range of topics, from the analytical treatment of various stress-strength models in a remarkably enlightening fashion by Martin Shooman, all the way to the application of probabilistic techniques to the design of an antenna by Mr. Moreno; from the design of materials strength tests to obtain increased information from smaller samples than conventional tests by the Drs. Heller, to the careful investigation of the variation in statistical distribution parameters with cycles to failure and stress to failure reported by Mischke and Wagner.

I would also like to draw to your attention, the very fine paper, Number 6B2, by Dr. Arthur Sorensen entitled, "A Statistical Analysis of Product Reliability due to Random Vibrations," to be presented in Session 6B, which presents a systematic development of a linear cumulative damage model for fatigue under random vibration.

As organizer of these two sessions, I hope that these summary thoughts will stimulate and encourage those workers in the field of mechanical reliability to publish their results and motivate others interested in advancing the state-of-the-art to proceed with the filling of the continuing serious gaps in our knowledge of basic strength data and mechanical-structural reliability technique areas.

I also would like to express my appreciation to Mr. Walter Gunkel, our Vice Chairman, and Dr. Robert Heller for reviewing the many submittals of papers for our two sessions.

RELIABILITY ESTIMATE OF A SPACE DEPLOYABLE ANTENNA

Frank J. Moreno
Radiation, A Division of Harris-Intertype
P. O. Box 37
Melbourne, Florida 32901

INDEX SERIAL NUMBER – 1065

Introduction

A reliability estimate of a Space Deployable Antenna (SDA) is calculated based on engineering design values and results of development testing. The main problem is one of estimating the reliability of a "one-shot" device which is primarily a mechanical system. The SDA is segmented in two main subsystems - an Upper Restraint Subsystem (URS) and a Mechanical Deployment Subsystem (MDS) - which were modeled separately then combined to arrive at an estimate of successful deployment probability.

Problem formulation and the modeling approach are heavily influenced by the design and development process associated with the system. As a consequence, the estimation techniques are constrained to utilize data that are made available at various stages of the process.

Principle of Operation

A number of rigid, parabolic shaped, ribs are mounted radially about a central hub and a mechanical deployment mechanism. In the stowed state, the rib tips are loaded in restraining devices where rib preload force is restrained at this point by a captivated cable. Upon command, a pair of cutting devices sever the cable. The preload force on each rib will then act to overcome the restraint force produced by the mechanical devices holding the rib tips. The ribs will "pop" free when the restraint force is overcome and deployment will proceed by means of a ball screw and carrier type of the mechanical deployment mechanism. The ribs are subsequently driven to deployment and latched in a fully deployed state by a set of torque motors in the MDS. An RF reflective surface is then provided by a flexible, metallic double mesh which is attached to the ribs and is pulled tight during the deployment process. Figure 1 is a picture of the deployment sequence of an antenna being deployed in a vacuum.

Model Formulation

The design and development of space flight hardware requires adherence to a set of stringent criteria. The SDA, in addition to being flight hardware, must be constructed such that it can survive for a prolonged period of time in earth orbit. Hence, the design requirements plus performance requirements (e.g., reflector tolerance - antenna gain loss due to reflector surface error) are tenable grounds for treating the antenna ribs as identical members. The notion of functionally independent ribs is also based on design detail - namely the ribs are individually connected to the MDS through mechanical linkages. The previous rationale is used as a basis for modeling the ribs as identical and independent members. The ribs in the stowed state are being constrained by the URS and upon initiation of the deployment sequence a force which is a combination of preload and torque motor force will act to pull the ribs loose and drive them to the fully deployed state. The URS will be overcome in the case where the forces acting on each rib are such that the combined preload and torque motor force is greater than the restraint force. Successful operation of the URS is therefore modeled from the perspective of forces acting on identical ribs. Although the SDA design is closely controlled, the forces are not known exactly. Hence, the freeing forces are treated as random variables and the restraint force as an unknown parameter. Developmental testing results tempered with engineering judgment are used to arrive at distributions which characterize the freeing forces. The MDS must therefore function properly and its torque motors must deliver an amount of applied force to each rib such that the antenna is fully deployed. Function and design testing of the MDS was conducted to determine what failure modes if any would prevail in a space environment. The tests consisted of activating the MDS a number of times and observing whether or not the mechanism went through the deployment cycle. The MDS was tested as a separate entity uncoupled from the ribs and remainder of the SDA. This particular MDS test plan was motivated by schedule, cost and development constraints.

The probability of successful deployment of the SDA can be written as

$$\Pr\{SDA=1\}=\Pr\{(CC=1)\wedge(MDS=1)\wedge(URS=1)\} \qquad (1)$$

where:
(URS=1) is the event indicating successful functioning of the URS
(MDS=1) is the event indicating successful functioning of the MDS
and:
(CC=1) is the event indicating the successful cutting of the restraining cable.

Rewriting (1)

$$\Pr\{SDA=1\}=\Pr\{CC=1\}\Pr\{(MDS=1)\mid(CC=1)\}\Pr\{(URS-1)\mid(MDS=1)\wedge(CC=1)\} \qquad (2)$$

The event (CC=1) by actual hardware design, is independent of the other events, hence (2) can be written as

$$\Pr\{SDA=1\}=\Pr\{CC=1\}\Pr\{MDS=1\}\Pr\{(URS=1)\mid(MDS=1)\wedge(CC=1)\}. \qquad (3)$$

The last term in (3) is the probability that the URS will function properly given that the MDS operates as designed and the restraining cable is cut.

Estimation of Deployment Probability

Probability of Success - Upper Restraint Subsystem

Success for the URS is achieved by freeing the ribs from the restraint and driving them to fully deployed state. The URS success is conditioned on the success of the MDS and CC as indicated in equation (3). There are two components of force action on each rib in such a manner as to overcome the restraint force (weight of the ribs and mesh is considered as part of the restraint force) - a force due to preload and a force due to the torque motors (these two forces are referred to as freeing forces). The torque motor force is applied as a consequence of the condition of MDS success. The forces due to preload and torque motors are not known with certainty and are treated as random variables. The distribution of these variables is suggested by design criteria/ information and the final "best" distribution is determined by engineering judgment. The parameters of the distribution of preload force are based on a graphical technique which uses design values associated with antenna rib deflection necessary for the given antenna design. A graph of preload force versus rib deflection (this graph is linear in the region of interest) was constructed which mapped the average value of rib deflection and three sigma design limits into the corresponding preload forces. Hence, an average preload force is associated with an average rib deflection and three sigma limits of preload force are associated with three sigma limits on rib deflection. Stringent design requirements leading to a "tightly" controlled antenna construction are advanced as tenable grounds for assuming a well behaved distribution of preload force (e.g., construction of the ribs is such that the deflection necessary to load them in the restraining mechanism is subject to small variations about the mean value). Based on the above engineering analysis and judgment, a normal distribution of the component of freeing force due to preload is hypothesized. The mean and standard deviation of the distribution of freeing force due to preload are calculated to be 5.15 lbs and 0.500 lbs respectively (3 sigma limits of ± 1.5 lbs). The direct measurement of the torque motor force as a component of freeing force acting on the ribs was not possible due to the separate testing of the subsystems (i.e., the MDS which houses the torque motors was tested separately and without the antenna ribs being attached). Hence, design values of average torque motor force and three sigma limits were used in establishing the average torque motor force at 2.75 lbs and the standard deviation at 0.275 lbs. The same criteria, stringent design requirements and controlled antenna construction, are appealed to for hypothesizing a normal distribution of freeing force due to the torque motors.

Let:
X_1=the amount of freeing force on the ith rib due to preload.
X_2=the amount of freeing force on the ith rib due to the torque motors.

where X_1 is normally distributed with mean μ_1 and standard deviation σ_1 (i.e., $X_1 \sim n(\mu_1,\sigma_1)$) and likewise $X_2 \sim n(\mu_2,\sigma_2)$.

The probability that the combined action of X_1 and X_2 will free and drive the ith rib to the deployed state is

$$P_i=Pr\{X_1+X_2>K\} \tag{4}$$

where K is the amount of restraint force that must be overcome. The distribution of the sum $X=X_1+X_2$ is required in order to compute the indicated probability. The distribution of X is the convolution of X_1 and X_2 which are independent and normally distributed random variables. The resultant distribution is also normal, i.e.,

$X\ n(\mu,\sigma)$
where:
$\mu=\mu_1+\mu_2$

$\sigma=(\sigma_1{}^2+\sigma_2{}^2)^{1/2}$

Equation (4) can now take the form

$$P_i=Pr\{X>K\}=1-Pr\{X\leq K\}. \tag{5}$$

Writing equation (5) in terms of the standard cumulative normal

$$P_i=(1-\overline{\Phi}(\frac{K-\mu}{\sigma}))_i\ . \tag{6}$$

The SDA is defined to consist of 12 identical ribs each of which is functionally independent. All 12 ribs must be released in order for the URS to successfully achieve its intended mission objective, hence, the probability of 12 ribs being released is simply

$$\prod_{i=1}^{12} P_i$$

where P_i is defined by equation (6). As a consequence of the model development $\prod_{i=1}^{12} P_i$ is the probability of URS success (conditioned on MDS and CC success). The last term of equation (3) is therefore equal to

$$Pr\{(URS=1)\,|\,(MDS-1)\wedge(CC=1)\}=\prod_{i=1}^{12} P_i$$

$$=\prod_{i=1}^{12}(1-\overline{\Phi}(\frac{K-\mu}{\sigma}))_i. \tag{7}$$

Probability of Success - Cable Cutting Devices

The cable cutting devices consist of a pair of guillotines in redundant configuration. The probability that the upper restraint cable is cut by means of the pair of guillotines is

$$PG=1-(1-pg)^2 \tag{8}$$

where:
pg=the probability that a single guillotine will successfully cut the cable.

Using the notation of equation (3)

$$\Pr CC=1-(1-pg)^2 \tag{9}$$

Calculation of Upper Restraint Subsystem and Cable Cutting Success Probability

The parameters of equation (7) must be determined in order to calculate the probability of URS success. The values of $\mu=7.90$ and $\sigma=0.5706$ were calculated for the normal distribution of $X=X_1+X_2$ based on design values of the parameters of the normal distributions on X_1 and X_2. The restraint force (K) is not readily quantifiable and as such is treated as an unknown parameter allowed to range over a set which cover the extreme values of restraint force. Setting $\Pr\{MDS=1\}=1$ in equation (3) and the incorporation of cable cutting success probability provides insight into the deployment probability as a function of restraint force (K). Equation (3), subsequently becomes

$$\Pr\{SDA=1\}=(1-(1-pg)^2)\left(\prod_{i=1}^{12}\left(1-\Phi\left(\frac{K-7.90}{0.5706}\right)\right)_i\right) \tag{10}$$

Values of $pg=0.99$ and $pg=0.999$ (used to bracket the estimate of single guillotine success probability) and $3.0\leq K\leq 6.0$ lbs. were inserted and the calculations carried out. Final results are displayed in graphic form in Figure 2. Inspection of Figure 2 indicates that $\Pr\{SDA=1\}>0.999$ for $K<5.67$ lbs. when $\Pr\{MDS=1\}=1$. Further design calculations indicate that a realistic limiting value of K is around 4 lbs. Hence, the URS success probability appears conservative. The estimation of MDS success probability will complete the computations necessary for the reliability model.

Probability of Success - Mechanical Deployment Subsystem

The results of tests conducted in the design and development phase were used to construct a lower bound on the probability of successful operation of the MDS. Succinctly, the MDS was cycled 400 times, under various conditions, to determine what failure modes, if any, would show up. The extensive testing did not produce any failures. The testing can be thought of as representing 400 Bernoulli trials during which 400 successes were observed.

Let:
p=probability of successful operation of the MDS on a single trial

then the maximum likelihood estimator for p is

$$\hat{p}=\frac{X_o}{n}$$

where:
X_o=number of successes
and:
n=total number of trials.

Using the results of the tests then

$$\hat{p}=1,$$

which says that the best point estimate for the true probability of success p is $\hat{p}=1$. A more revealing statistic at this point is the lower bound on the true probability of success (p). The arguments leading up to and the development of the following lower bound can be found in References (1) and (2). A 95 percent Lower Bound (LB) on the parameter p is given by the following:

$$LB=\frac{X_o}{X_o+(n-X_o+1)F_{0.95}(2(n-X_o+1),2X_o)} \tag{11}$$

where $F_{0.95}(2(n-X_o+1),2X_o)$ is random variable with the variance ratio distribution and is a function of two parameters (degrees of freedom).

Substituting in equation (11) for $X_o=400$ and $n=400$ and using tables for the cumulative F distribution[2]

$$LB=0.993.$$

Based on the test results we are 95 percent sure that the true probability of successful operation of the MDS is no smaller than 0.993.

Probability of Success - Space Deployable Antenna

The estimation of equation (3) consists of using the lower 95 percent statistical bound for MDS success and a lower bound derived by design considerations on the URS. Making the appropriate substitutions in equation (3) yields

$$\Pr\{SDA=1\}\simeq(0.999)(0.993)=0.992.$$

This estimate is compatible with the initial design goal of 0.99 for probability of successful deployment.

Conclusion

Design and test data can be used to provide an efficient and economical base for reliability estimation.

References

(1) Hald,A., 1952, Statistical Theory with Engineering Applications, John Wiley and Sons, New York

(2) Brownlee,K.A., 1960, Statistical Theory and Methodology in Science and Engineering, John Wiley and Sons, New York

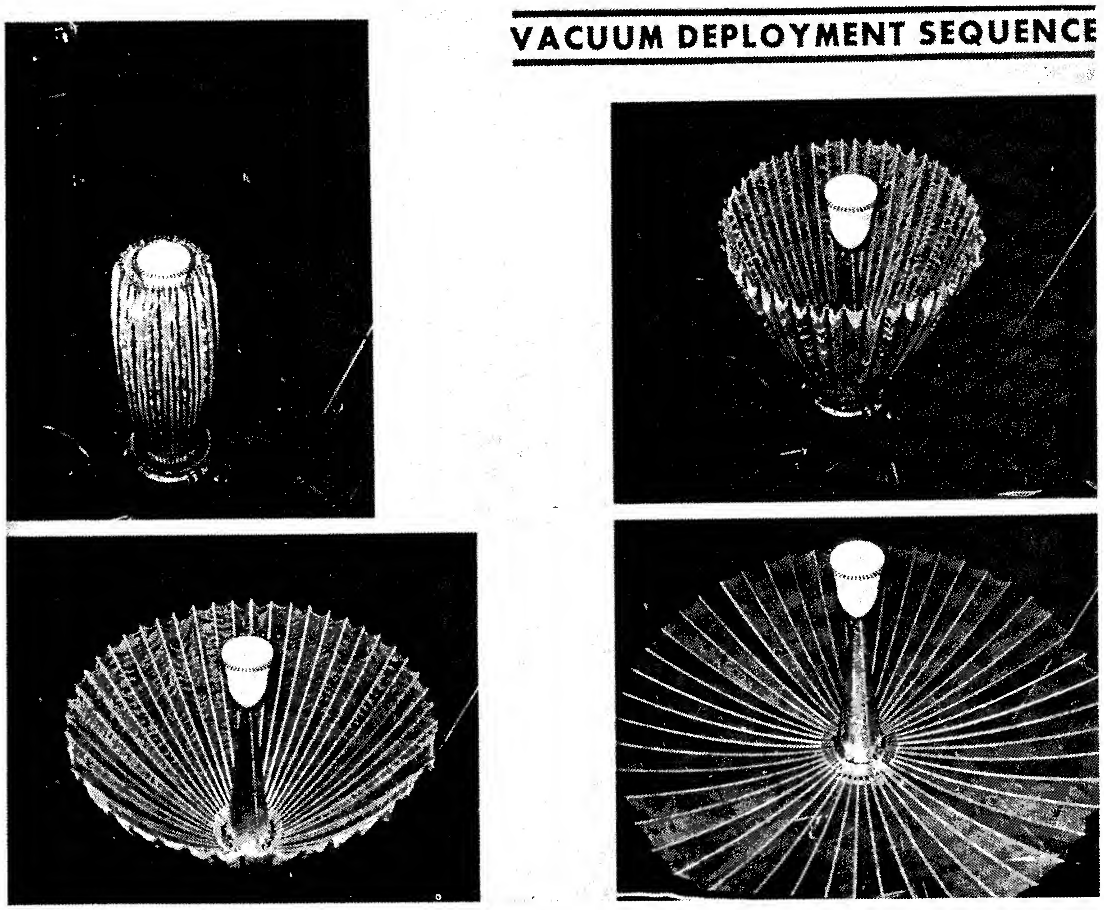


Figure 1. Antenna Being Deployed in a Vacuum

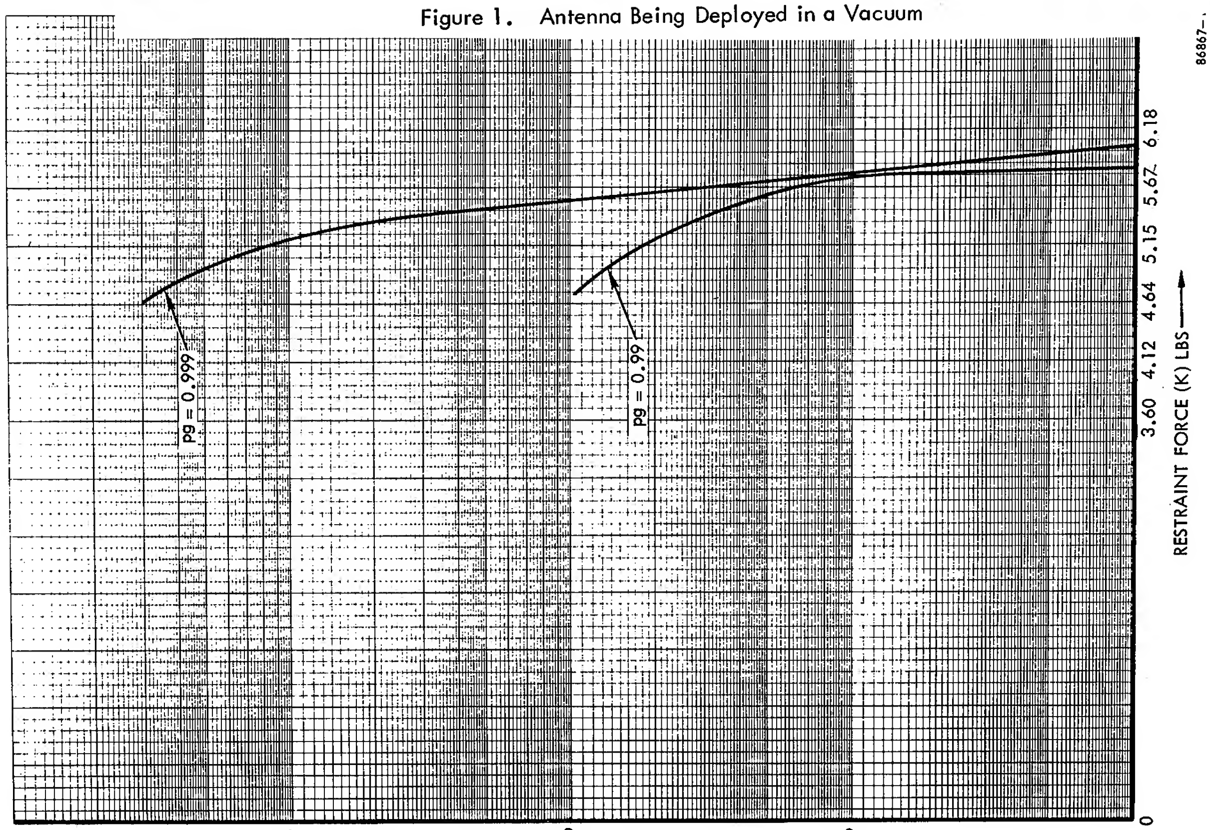


Figure 2. Probability of Antenna Deployment Versus Restraint Force Assuming Pr {MDS = 1} = 1

TIME DEPENDENT STRESS-STRENGTH MODELS
FOR NON-ELECTRICAL AND ELECTRICAL SYSTEMS

Part I of II

Leonard Shaw, + Martin Shooman, + Robert Schatz*

ABSTRACT

Structural modeling of systems and failure rate modeling of components has proved quite successful in the reliability analysis of electrical systems. In the case of mechanical, pneumatic, hydraulic, etc. systems, such techniques are hard to apply because of the distributed nature of the devices. In such cases stress-strength models are often used. These models are good for single stress situations, but must be modified to describe the commonly occurring situation of repeated stresses.

A general reliability modeling technique is developed for time and/or cyclic dependence of the stress and strength distributions use to characterize failure modes in the stress-strength interference method. These distributions may change due to aging and cumulative damage, or due to the information gained by a history of non-failures.

Related earlier work by Freudenthal,[1] Reethof[2] and Shooman[3] has been generalized and unified. By focusing on the degree of knowledge about the particular stress and strength involved one is able to define 9 different models, several of which model important practical cases. Since time and/or cycle dependence of the stress is included, the model results in a time (and/or cycle) dependent reliability function and an equivalent failure rate. This permits an analyist to devide his system into a lumped electrical (and mechanical) portion and a distributed portion. Conventional failure rate modeling can be used to derive a lumped reliability function and stress-strength modeling to obtain a distributed reliability function. The product of these two functions yields the system reliability function.

Introduction

Reliability analysis of a system usually comprises two steps. Failure descriptions for individual components or subsystems are determined by statistical analysis of test data, and then structural relations among the components in the system are formulated to relate the system reliability to the component or subsystem reliability. Many components especially electrical ones, are amenable to tests from which failure rates (hazard functions) can be estimated. However, the failure mechanisms for structural, hydraulic and pneumatic components, among others, are not well described by average failure rates because the failures are generally caused by isolated, identifiable stresses. In addition, these system elements are generally distributed (rather then lumped), costly, more often custom (rather then stock) designs, and in general difficult to define, isolate and test. Effective failure rates can be defined for components in the latter category by using probability descriptions for the stress and strength to determine the probability of failure after a single stress applications, and then combining that probability with a model for the times at which stresses occur. The single-stress failure probability is computed from the stress and strength distributions by the stress-strength-interference (SSI) technique.[1] The times at which stresses occur may be cyclical[2] or random.[3] The result is, in either case, a reliability function where the equivalent failure rate is a function of strength distribution, and stress occurrence law parameters.

Although the motivation for developing stress-strength-time (SST) models is to describe the reliability of non-electrical elements it can also be used for electrical elements. One example of such a case is the attempt to describe capacitor failures in terms of a more microscopic view point (sometimes called reliability physics). One might view the dielectric breakdown voltage as a strength, and the applied voltage as a cyclic or random stress. Clearly the problem is now described in terms of SST theory.

This paper generalizes and unifies the work in reference 1,2 and 3, as motivated by examples in reference 4, with a view toward getting exact and approximate reliability expressions for mechanical components operating in environments with repeated stresses. Time variations in stresses (due to operating sequences) and in strengths (due to aging or secondary stresses) are also considered.

Each of the component variables, stress or strength, can be classified in one of the following three levels of uncertainty.

1. Known Stress or Strength - The variable is either constant or varies in some known predictable manner. If both stress and strength are known, failure modeling is deterministic rather than probabilistic. The device succeeds if strength > stress, and fails if strength ≤ stress. Of course known implies that the manufacturing process is well controlled so the parameters are predictable or a simple nondestructive test is available to determine the parameters.

2. Random-fixed (stress or strength) - The variable is either constant or varies in time in a known manner; however, the constants of the model are unknown. It is assumed that enough data has been recorded in the past so that a probability density function for the strength (or stress) is known. It is assumed that any test to determine the variable precisely (as in 1. above) is too costly or destructive. Since the variable is a fixed (or predictably changing) function, after each success we become more certain that the unknown strength is high. Therefore, this situation calls for a dependent probability calculation.

3. Random - Independent stresses (or strengths) Not only is a single stress value sufficiently

+Both with the Department of Electrical Engineering/Electrophysics of the Polytechnic Institute of Brooklyn.

*Westinghouse Advanced Reactor Division - Formerly Astronuclear Laboratory
This work was in part supported by the Westinghouse Astronuclear Laboratory under consulting agreement 59-CYO-20931 and in part by the office of Naval Research, Statistics and Probability Program under contract N00014-67-A-0438-0013.

unknown so that it is well described as a random variable, but successive stresses are so unrelated as to be statistically independent. Observation of one stress level gives no information about the size of subsequent stresses. Although probably of little practical interest, we could also consider random, independent strengths, say due to independent fluctuations of a third influence, like temperature, on a mechanical failure mode.

Examples of failure modes suitable for SST calculations are tensile, creep and compressive forces vs. tensile strength, creep resistance, and buckling resistence, respectively, as well as temperature stress over a time interval vs. corrosion resistance, and reactivity demand on a nuclear fuel element vs. available reactivity. Although a stress-induced failure or a successfully resisted stress occurrence will take place over a non-zero time interval, these events can be assumed to be instantaneous if a much longer operating time scale is of interest for the system.

1. Cyclic Occurrences - These are changes in the stress which follow a cyclic pattern. These may be due to natural seasonal changes or day/night temperature cycles or man made on-off, up-down, etc. cycles. We include in this category cyclic changes which have a fixed period and those with a non-constant time between cycles. The reliability is a function of the number of cycles, n, rather than time. Examples are the operating cycles of a relay or the number of take offs and landings of an aircraft.

2. Random occurrences - This case refers to situations in which the times between stress occurrences are random rather than known. If we assume, for example, that stresses occur infrequently and that each stress occurrence is independent of all others, we have a Poisson probability law for stress occurrence. Random occurrences include not only the Poisson probability law but other occurrence laws as well.

Stresses and strengths may vary during a long operating interval, in relation to the passage of time or to the number and/or severity of previous stresses. We will use the following classification for such time variations.

A. Aging - Aging describes changes with time in the parameters of the model. Most commonly this is modeled as a shift of the mean and/or variance, e.g., a linear decrease in strength. A simple examply is the corrosion in a liquid cooling system.

B. Cyclic Damage - An item may experience a change in its strength as the device undergoes repeated operating cycles. Thus, the strength density is a function of the number of cycles n. An example of this phenomena is the shortened life of a light bulb which is subject to many on-off cycles rather than allowed to burn continuously.

C. Cumulative Damage - A device is said to suffer cumulative damage when its decrease in strength is determined by the size as well as the number of previous stresses. An example would be air leakage from a space craft due to meteorite collisions puncturing the skin. We assume that larger meteorites make larger holes creating more air leakage.

Stress-Strength Interference Theory

Before we begin to broaden stress-strength interference, (SSI) theory to include time dependence we will briefly summarize conventional stress-strength theory. Clearly the concept of a part strength surviving an applied stress fits in well with some of the design philosophy of Mechanical and Civil Engineering. Considerable work has been done in the past on SSI theory. (reference 5). If we deal with a single strength, y, and a single stress x, assume these are random variables with density functions f(x) and g(y), then we can formulate our problem in two equivalent ways. The direct approach is to let the random variable w, represent the excess of strength over stress

$$w = y - x \qquad (2\text{-}1)$$

The probability density for w is denoted u(w). Clearly the part succeeds when w>0 one fails when $w \leq 0$. Thus, the probabilities of success and failure are.

$$p_s = P(w>0) = \int_0^{\infty} u(w)\, dw \qquad (2\text{-}2)$$

$$p_f = 1 - p_s = P(w \leq 0) = \int_{-\infty}^{0} u(w)\, dw \qquad (2\text{-}3)$$

One must take care to formulate the problem correctly in the case where failure can occur symetrically for both positive and negative stresses. For example, a cantalever beam may break due to a negative load (downward) or positive load (upward). In fact one might even have different failure mechanisms yielding different equations as in the case of a beam failing under tension and compression.

Sometimes the SSI problem is formulated by letting v=y/x and defining success as P(v>1). This approach requires careful handling of areas if x and y take on negative values. In the special case of symmetrical failures it may be more convenient to work with v if one requires that the sign of the stress and strength variables always be the same.

In essence equations 1, 2, and 3 reduce stress-strength theory to a transformation of random variables to compute u (w), followed by an integration. Since the transformation is a difference (sign change and sum) the change of variables yields the convolution integral. (see ref. 3,p.78). In the case where f(x) and g(y) are Gaussian densities, the transformed variable u(w) is Gaussian and Eqs. (2) and (3) are evaluated by looking in a normal probability table.

In the cases where a normal distribution is a poor model for x and/or y one can use Rayleigh, Weibull, beta, etc. distributions. In such a case, the transformation of random variables is generally not tractable and one resorts to numerical approximation and tabulation of results (ref. 5) or analytical approximation of the integrals involved. Another simple approach is to compute the moments of w ignoring the distributions of x and y.

$$E(w) = E(x) - E(y) \qquad (2\text{-}4)$$

$$Var(w) = Var(x) + Var(y) \qquad (2\text{-}5)$$

In the case where we know the distribution of w (if Gaussian, the normal table, otherwise we must consult the literature (ref. 3,p.394) or generate our own table.) Eqs. 4 and 5 allow us to enter the appropriate probability table. Even if we do not know the

distribution of w, we can use Tschebyschev's or Gauss's Inequality to bound the result.

The alternate formulation of the probability of success equation is to formulate a joint density function of x and y, $\phi(x,y)$, and compute the area under this function where $y<x$. Assuming independence and non-negative domains for x and y.

$$\phi(x,y)= f(x)\ g(y) \tag{2-6}$$

$$p_s= \int_0^\infty [\int_0^{x=y} f(x)g(y)\ dx]\ dy = \int_0^\infty g(y)\ [\int^{x=y} f(x)dx]\ dy \tag{2-7}$$

The results are identical to those given by Eqs. 2-1 2-2, 2-3.

If we are to use SSI theory for system design, it is important to know how p_s changes as we change the densities of y and x. The simplest approach is to assume Gaussian densities for x and y, compute p_s and the partial derivatives of p_s with respect to the means and variances of x and y. Of course the results will only be exact for Gaussian densities but should still be indicative of the trends to be expected in the general case. Denoting the means and standard deviations of x and y by μ_x, μ_y, σ_x, σ_y, expanding p_s in a multivariable Taylor series about an operating point and truncating higher order terms yields

$$p_s(\mu_x, \mu_y,\sigma_x,\sigma_y)=p_{s_o}+\frac{\partial P_s}{\partial \mu_x}\Delta\mu_x+\frac{\partial P_s}{\partial \mu_y}\Delta\mu_y \tag{2-8}$$

$$+\frac{\partial p_s}{\partial \sigma_x}\Delta\sigma_x+\frac{\partial p_s}{\partial \sigma_z}\Delta\sigma_y$$

In the case of normal distributions using Eqs. 2-1 2-2, 2-4 and 2-5,

$$pf=\int_{-\infty}^{\alpha} \frac{e^{-t^2/2}}{\sqrt{2\pi}}\ dt \tag{2-9}$$

$$\alpha=\frac{\mu_y-\mu_x}{\sqrt{\sigma_x^2+\sigma_y^2}}$$

Differentiating p_f with respect to the four parameters of α the upper limit of the integral, yields

$$\frac{\partial p}{\partial \mu_y} = \frac{e^{-\alpha^2/2}}{\sqrt{2\pi}}\ \frac{1}{\sqrt{\sigma_x^2+\sigma_y^2}} = u(0) \tag{2-10}$$

$$\frac{\partial p}{\partial \mu_z} = -u(0) \tag{2-11}$$

$$\frac{\partial p}{\partial \sigma_x} = -u(0)\frac{\mu_w}{\sigma_w^2}\sigma_x \tag{2-12}$$

$$\frac{\partial p}{\partial \sigma_y} = -\mu(0)\frac{\mu_w}{\sigma_w^2}\sigma_y \tag{2-13}$$

Clearly, the partials with respect to μ_x and μy are equal in magnitude and opposite in sign. Also examining Eqs. 2-12 and 2-13, we see that the success probability is more sensitive to changes in the larger standard deviation. The relative magnitudes of variation due to changes in the μ's and σ's is best evaluated by numerical substitution into Eqs. 2-10 2-11, 2-12, 2-13.

Stress Strength Time-Cyclic Repetitions

Cyclic stress repetitions either occur at known times; or correspond to situation in which the order of stress occurrences is important, but not the times of occurrence. In such cases the reliability function R (t) becomes R_n, with a discrete argument, representing the probability that n successive stresses do not cause a failure. If the occurrence times t_n are known, a continuous-time reliability R(t) can also be defined by

$$R(t) = R_n;\quad t_n<t<t_{n+1} \tag{3-1}$$

as shown in Figure 3-1. It is important to note that if the failure mechanism remains fixed while a differset of occurrence times is chosen, then the new R(t) is found by simply shifting each discontinuity in Fig. 3-1 to the corresponding new t_n (i.e. a distortion of the abscissa). Thus, R(t) is essentially the same for periodic or non-periodic known stress occurrence times.

It will also be useful to think of a discrete survival probability function $R_{n,n-1}$ the probability of success on occurrence n, given that the (n-1) preceding occurrences have been successful. Calculations of the discrete R_n and $R_{n,n-1}$ will now be outlined for the nine possible combinations of known, random - fixed and random-independent stresses and strengths. The number of derivations can be reduced by eliminating details for symmetrical cases in which stress and strength reverse roles. We begin with the three cases in which both variables are in the same category. Stress categories will be indicated by numbers and strength categories by letters (see Figure 3-20

Case 1.a. Known Stress and Strength - This trivial case is of little interest except, perhaps, when aging takes place. The probability of success on the i^{th} occurrence is

$$p_{s_i} = \begin{cases} 0, & y(t_i)\leq x(t_i) \\ 1, & y(t_i) > x(t_i) \end{cases} \tag{3-2}$$

if x and y are constants,

$$R(t) = P_{s_i}\quad \text{for } t > t_1 \tag{3-3}$$

for time varying variables

$$\begin{aligned} R(t) &= 0 \text{ if } y(t_i)\leq x(t_i) \text{ for some } t_i\leq t \\ &= 1 \text{ if } y(t_i) > x(t_i) \text{ for all } t_i\leq t \end{aligned} \tag{3-4}$$

As a special case of (3.4), if stress is constant or increasing, and strength is decreasing then the discrete reliability is

$$R_n(t_n) = \begin{cases} 0, & y(t_n)\leq x(t_n) \\ 1, & y(t_n)> x(t_n) \end{cases} \tag{3-5}$$

and R(t) depends on the strength at the most recent occurrence.

Case 2.b. Random-Fixed Stress and Strengths.

The first occurrence of a random-fixed stress on a part with random-fixed strength is precisely the situation analyzed in Section 2 by the SSI approach to get p_{s_1}. If both variables remain constant during repeated stresses, success or failure on the occurrence determine success or failure on all future ones. Thus

$$R_n = P(s_1, s_2, \dots s_n) = p_{s_1} \tag{3-6}$$

The reliability can depend on time in this case if the variables have some deterministic time (or cycle) dependence, rather than being constant in time. If, for example, the fixed-random strength has a degrading mean which decreases with time, expressible by the relation

$$\begin{aligned} y(t) &= y(0) - a(t) \\ a(0) &= 0 \\ \frac{da}{dt} &\geq 0 \ ; \quad t > 0 \end{aligned} \tag{3-7}$$

then the success probability at time t_n can be expressed variously as

$$\begin{aligned} p_{s_n} &= P[\, y(t_n) > x(t_n)] \, - \\ &= P[\, y(0) > x(0) + a(t)] \\ &= P[\, w(0) > a(t)] \end{aligned} \tag{3-8}$$

The last expression can be evaluated using the density function u(w) described above. The resulting reliability would be

$$R_n = p_{s_n} = \int_{a(t_n)}^{\infty} u(w)\, dw \tag{3-9}$$

for this case of random - constant stress and random deterministically decreasing strength. The discrete survival function which is the conditional probability of surviving the n'th stress given that you have survived all of the previous n-1 stresses.

$$R_{n,\, n-1} = P[\, s_n \ [s_1, s_2 \dots s_{n-1}] \tag{3-10}$$

This survival function is related to reliabilities at t_n and t_{n-1} by

$$R_n = R_{n-1} \cdot R_{n,n-1} \tag{3-11}$$

In this case the numerical function can be computed from (3.11) and (3.9).

Case 3.c. Random - Independent Stress and Strength

In this case each stress occurrence is a random event, independent of all others. A strength that varies randomly and independently at successive stress times is hard to imagine physically, although it could occur through the action of a secondary random stress which weakens the parts resistance to the primary stress, e.g. a reverse change in strength proportional to a randomly varying temperature.

This situation corresponds to a constant survival

$$\begin{aligned} R_{n,\, n-1} &= P[\, s_n \ [s_1, s_2, \dots s_{n-1}] \\ &= p_{s_i} = P[\, w_1 > 0] \end{aligned} \tag{3-12}$$

which can be computed via SSI methods. The corresponding discrete reliability will be

$$R_n = p_{s_i} \cdot R_{2,1} \cdot R_{3,2} \cdots R_{n,\, n-1} \cdot = (p_{s_1})^n . \tag{3-13}$$

If aging or cyclic time variations occur in the stress or strength distributions, the $R_{k,k-1}$ in (3-12) and (3-13) can be replaced by

$$R_{k,\, k-1} = P[\, w_k > 0] \tag{3-14}$$

when u (w_k) is defined using appropriate densities $f(x;t_k)$ and $g(y;\ t_k)$.

Four more cases, 1b, 2a, 1c, and 3a can be described as special cases of the three just considered, as suggested by the arrows in Figure 3-2. For example, 2b becomes 2a when the strength density is made an impulse, corresponding to a known strength. Similar use of impulse densities allows the previous discussion to include cases 1b, 1c and 3a.

The remaining distinct cases are the symmetrical 3b and 2c, of which we will examine the more practical former one.

Case 3.b. Random - Independent Stress, Random - Fixed Strength.

Even in the simplest version of this case, with a constant random strength and identically distributed stresses the survival function is not constant. If a part with unknown strength survives one stress, it is more likely to have a high strength than a low strength. This means that the conditional density of the strength after (n-1) successes will be a function $g_n(y)$ depending on n

$$g_n(y)\, dy = P[\, y < \underline{y} \leq y + dy \ [\underline{x}_i < \underline{y};\ i = 1, 2, \dots (n-1)] \tag{3-15}$$

with $g_1(y)$ representing the initial strength density.

Clearly, $R_{n,n-1}$, the probability that $x_n < y$ given successes on all previous stresses is, as in (2.7),

$$R_{n,\, n-1} = \int_{y=0}^{\infty} \int_{x=0}^{y} f(x) g_n(y)\, dx\, dy. \tag{3-16}$$

Furthermore, the basic conditional probability definition leads to the relation

$$R_{n+1} = R_n \cdot R_{n,\, n-1} \tag{3-17}$$

between successive reliabilities. This recursion is started with R_1 computed as $R_{1,0}$ from (3.16)

The appendix shows how the proceding three equations can be combined to get the following general expressions for the cycle dependent strength density and the survival.

$$g_n(y) = g_1(y) \left[\int_0^y f(x)\,dx \right]^{n-1} \Big/ R_{n-1} \quad n \geq 1 \qquad (3\text{-}18)$$

$$R_{n+1,n} = \left[\int_{y=0}^{\infty} g_1(y) \left(\int_0^y f(x)dx \right)^{n+1} dy \right] \Big/ R_n \qquad (3\text{-}19)$$

The following simple example shows how to use and interpret these results.

Figure 3.3 shows rectangular densities for f(x) and $g_1(y)$, which have been chosen for computational simplicity. The first steps are to compute R_1 and $R_{2,1}$ from

$$R_1 = \int_{y=0}^{\infty} g_1(y) \int_{x=0}^{y} f(x)\,dx\,dy \qquad (3\text{-}20)$$

$$g_2(y) = g_1(y) \int_{x=0}^{y} f(x)\,dx \,/\, R_1 \qquad (3\text{-}21)$$

Both of these terms require computation of the integral $\int_0^y f(x)\,dx$ of the rectangular density, resulting in the triangular function of y shown in Figure 3.4. Multiplication of that function by the $g_1(y)$ in Fig. 3.3 produces the numerator of (3.21) as shown in Figure 3.5. The R_1 of (3.20) is the area under the curve of that figure, which turns out to be unity minus the area A of the shaded triangle in that figure, which turns out to be unity minus the area A of the shaded triangle in that figure

$$A = 1-R_1 = \frac{1}{2}\left[\frac{1}{y_u-y_\ell}\left(1-\frac{y_\ell-x_\ell}{x_u-x_\ell}\right)\right](x_u-y_\ell)$$

$$= \frac{(x_u-y_\ell)^2}{2(y_u-y_\ell)(x_u-x_\ell)} \qquad (3\text{-}22)$$

Finally, the $g_2(y)$ in Figure 3.6 is simply a scaled version of the numerator curve in 3.5, due to division by R_1 in (3.21). The change in the strength density functions from $g_1(y)$ to $g_2(y)$ is worthy of note. That portion which contains the high stress region of stress density function f (x) shown specified decreases monotonically in area as cycle number n increases effectively truncating $g_n(y)$.

The effect of this whole computation has been to change the shape of the $g_n(y)$ function to reflect the fact that the item being stressed is more likely to be one of the high strength members of the population described by the original density $g_1(y)$. The effect has been called "tail erosion" by Reetoff.[2] If we were to carry the process through again to find $g_3(y)$ we would obtain a density function with the shape shown in Fig. 3.7. The progression of the " tail erosion" process can be clearly seen. As the "erosion" process continues, we asymptotically approach a rectangular density as $n\to\infty$ when n is large for $g_n(y)$ extending between $y = x_u$ and $y = y_u$. In other words, all the overlap area between x and y has been eroded and R_n does not change as n increases further. Of course if the density functions are not truncated as in this example, the same effect goes on but it is not as easy to describe the asymptotic result. If the unknown strength is a function of time, the time variation is carried along in the above process as a parameter.

It is worthwhile to consider the difference between the reliability computed as outlined here, and a more naive or approximate approach which assumes that all survivals are the same as the first; i.e. $R_{n,n-1}=R_1$. This amounts to using case $3c_4$ to approximate 3b. An example taken from reference [4], has gaussian stress and strength densities with standard deviations

$$\sigma_x = 5420 \text{ kpsi} \qquad (3\text{-}23)$$

$$\sigma_y = 6710 \text{ kpsi}$$

and a difference of means (3-24)

$$\mu_y - \mu_x = 43,000 \text{ kpsi}$$

Computations using equations 3-18 and 3-19 show that

$$R_{60} = 0.999981 \qquad (3\text{-}25)$$

while the conventional constant failure rate approximation gives

$$\hat{R}_{60} = (R_1)^{60} = 0.999974 \qquad (3\text{-}26)$$

A comparison of results (3-25 and (3-26) indicates that SST methods yield more accurate estimates of reliability for certain types of components, and that this estimate can be <u>higher</u> than would be computed using more conventional techniques.

Another way of looking at this is that

$$R_{60} \approx (R_1)^{44} \qquad (3\text{-}27)$$

Rather than $\hat{R}_{60}$ of equation (3-26).

<u>Reliability for Random Stress Occurence Times</u>

It is possible to write the following general expression for the reliability function R(t) when the stress occurrence times are random

$$R(t) = \Pi_0(t) R_0 + \Pi_1(t) R_1 + \ldots + \Pi_n R_n \ldots \qquad (4.1)$$

where $\Pi_n(t)$ is the probability of n occurrences in the 0 to t times interval, and R_n is the probability of n successes, described above. (Even the known occurrence time case is covered by this expression. In that case, exactly one $\Pi_n(t) = 1$ and all the rest are zero, for each value of t.)

The R_n functions of Sec. 3 must be combined with a probabilistic model for occurrence times in order to evaluate (or approximate) the infinite sum in (4-1). The actual application must be used to guide the choice of an occurrence model. Some models have the advantages of being physically reasonable in a wide variety of situations, as well as being mathematically tractable.

It is often reasonable to assume that the occurrence times, after a reference time t_r, are independent of the actual occurrence times before t_r, but similarly distributed over corresponding time intervals of equal width. In addition, it is often reasonable to assume that in a small interval (Δt), at most one occurrence can take place with an occurrence probability α(Δt) proportional to the width of the small interval. It can be shown [5] that these

mild assumptions require that the number of occurrences N_t in the interval 0 to t governed by the Poisson probability law

$$P[N_t = k] = \frac{e^{-\alpha t}(\alpha t)^t}{k!} \tag{4.2}$$

$$= \pi_k(t)$$

Furthermore, it follows that in this Poisson occurrence case, the times θ_i between successive occurrences

$$\Theta_i = t_i - t_{i-1} \tag{4.3}$$

are independent, identically distributed random variables with exponential density functions

$$f_{\theta_i}(\theta) = \alpha e^{-\alpha\theta} \; ; \qquad \theta \geq 0 \tag{4.4}$$
$$= 0 \text{ otherwise}$$

This property suggests an approach to forming other, non-Poisson occurrence models. These related models assume that the inter-occurrence times θ_i are independent, but replace (4.4) by some other density function (which need only be zero for negative θ). Unfortunately, these similar models do not have such simple expressions as (4.2) for $\Pi_k(t)$, which is needed to evaluate (4.1). (Another generalization is the generalized Poisson law in which the short time probability of a single occurrence in Δt becomes $\alpha(t)\,\Delta t$.)

A further justification for Poisson occurrence models is noteworthy. If, for example, a stress occurrence corresponds to a trajectory-correcting thrust, then this takes place when a control system decides that a velocity error function has reached an intolerable level. The velocity error may well be represented by a Gaussian random process, since the Central Limit Theorem would predict this model if velocity errors are caused by many independent accelerations, say due to meteorites. Thus in this case, stress occurrences correspond to the times that a Gaussian process rises above a control decision threshold (see Fig. 4-1).

It is known that such up-crossings of a Gaussian process are, indeed, approximately described by the Poisson occurrence law of (4.2).[6]

The remainder of this section will be devoted to the case of Poisson occurrences, but it should be emphasized that computational evaluation of (4.1) for other occurrence laws, as determined from data, is quite feasible.

The following reliability calculations combine the Poisson occurrence law of (4.1) (in which α is the mean number of occurrences per unit time) with the various failure cases in Sec. 3.

Case 1-a; Constant, known stress x_0 and strength y_0:

$$R_n = 1 \text{ if } x_0 < y_0$$

$$= 0 \text{ if } x_0 \geq y_0, \; n \geq 1$$

Therefore

$$R(t) = \sum_{k=0}^{\infty} 1 \cdot P[N_t = k] =$$

$$= P[\underline{\text{any}} \text{ number of occurrences}]$$

and

$$R(t) = 1 \text{ for } x_0 < y_0 \tag{4.5}$$

$$R(t) = P[N_t = 0] \; ; \; x_0 \geq y_0$$

$$R(t) = e^{-\alpha t}; \qquad x_0 \geq y_0 \tag{4.6}$$

Notice that when $x_0 \geq y_0$ (4.6) corresponds to a constant hazard with $\lambda = \alpha$

Case 3a: Constant known strength, random independent stresses at each occurrence.

$$R_n = (p_{s_1})^n$$

$$R(t) = e^{-\alpha t}\left[p^0_{s_1} + p_{s_1}\,\alpha t + p^2_{s_1}\,\frac{(\alpha t)^2}{2} \cdots\right]$$

$$= e^{-\alpha t}\, e^{\alpha p_{s_1} t}$$

$$R(t) = e^{-\alpha(1-p_{s_1})t} \tag{4.7}$$

Here, too, there is a constant hazard of

$$\lambda = \alpha(1-p_{s_1})$$
$$= \alpha\, p_f$$

i.e. the occurrence rate times the single-occurrence failure probability.

Case 1-b: Constant, known stress; random fixed strength.

$$R_n = p_{s_1} \qquad n = 1, 2 \ldots 1$$

$$R(t) = e^{-\alpha t}\left[1 - p_{s_1}\,\alpha t + p_{s_1}\,\frac{(\alpha t)^2}{2} + \ldots\right]$$

$$= e^{-\alpha t}\left[(1 - p_{s_1}) + (p_{s_1} + p_{s_1}\,\alpha t + \ldots)\right]$$

$$= e^{-\alpha t}(1-p_{s_1}) + e^{-\alpha t}\, p_{s_1}\, e^{\alpha t}$$

$$R(t) = p_{s_1} + (1-p_{s_1})\, e^{-\alpha t} \tag{4.8}$$

Although the hazard is not constant here, it can be computed from the basic definition

$$z(t) = -R'(t)/R(t)$$

$$z(t) = \frac{\alpha\ (1-p_{s_1})e^{-\alpha t}}{(1-p_{s_1})e^{-\alpha t} + p_{s_1}}$$

$$z(t) = \frac{\alpha}{1 + e^{\alpha t}\ p_{s_1} / (1-p_{s_1})} \tag{4.9}$$

Figure 4.2 shows this hazard function for several values of p_{s_1}. Note that the hazard is constant at $\lambda = \alpha$ when $p_{s_1} = 0$, and is approximately exponential

$$\lambda \approx \alpha(1-p_{s_1})e^{-\alpha t}$$

when p_{s_1} is nearly equal to 1.

Finally, the most interesting case of Poisson occurrences corresponds to random independent stresses and a fixed, but randomly distributed strength. Substitution of the R_n computed in Case 3-b of Sec. 3 into (4.1) yields a general expression for R(t) which, in most cases would be best evaluated computationally. The precise result can be bounded by noting that for any stress and strength distributions,

$$(p_{s_1})^n \leq R_n \leq p_{s_1} \tag{4.10}$$

These inequalities express the facts that the probability of n successes must be no greater than the probability of one success; and that the conditional information of previous successes can only increase the probability of subsequent successes.

Combination of (4.1) and (4.10) gives the following general bounds on R(t)

$$e^{-\alpha(1-p_{s_1})t} \leq R(t) \leq p_{s_1} + (1-p_{s_1})e^{-\alpha t} \tag{4.11}$$

where we have used the fact that the upper and lower bounds on R_n correspond, respectively, to Cases 1-b and 3-a.

A more explicit expression for R(t) follows from the general R_n. Substitution of (3-19) and the Poisson occurrence probabilities into (4.1) produces

$$R(t) = e^{-\alpha t} \sum_0^{\infty} \frac{(\alpha t)^n}{n!} \int_0^{\infty} g_1(y) \left[\int_0^y f(x)dx \right]^n dy \tag{4.12}$$

Changing the order of integration and summation leads to

$$R(t) = e^{-\alpha t} \int_0^{\infty} g_1(y) \sum_0^{\infty} \frac{\left[\alpha t \int_0^y f(x)dx\right]^n}{n!} dy$$

The final expression follows from identifying the sum as an exponential and writing the x-integral in terms of the distribution function

$$F_x(y) = \int_0^y f(x)dx, = 1 - \bar{F}_x(y)$$

namely

$$R(t) = \int_0^{\infty} g_1(y)\ e^{-\alpha t\, \bar{F}_x(y)}\ dy \tag{4.13}$$

The corresponding hazard is in

$$z(t) = \frac{\alpha \int_0^{\infty} g_1(y)\, \bar{F}_x(y)\ e^{-\alpha t\, \bar{F}_x(y)}\ dy}{\int_0^{y} g_1(y)\, e^{-\alpha t\, \bar{F}_x(y)}\ dy} \tag{4.14}$$

Computational evaluation of these general expressions for R(t) and z(t) is straightforward once stress and strength densities have been prescribed. Approximate expressions for the case of small (αt) (infrequent stress occurrences and/or short time operation) follow from a corresponding approximation of (4.12) by the first few terms in the summation.

It is germain at this point to digress to discuss one potentially important use of the model for R(t) bounded by Eq. 4.11. We see that from Eqs. (4.7), (4.8) and (4.9), for certain conditions, we have an exponentially decreasing hazard. Although many electronic components have been found to have a constant hazard, two notable exceptions immediately come to mind. Integrated circuits and capacitor are known to have a decreasing hazard. It is not hard to imagine a microscopic failure model for a capacitor involving the dielectric strength and the applied voltage stress. One might be able to construct a similar SST model for an integrated circuit. Thus, the models developed here may be important in the study of the microscopic failure behavior of devices which is often called reliability physics.

Reliability with Aging or Cyclic Damage

The calculations in the previous section assumed that the probability densities for random stresses and strengths were not explicitly time-dependent. In many practical situations these densities will not remain fixed. If they change with the passage of time, the effect is called aging (generally a decrease in strength); whereas, if the changes correspond to the number of stress occurrences, the effect is called cyclic damage. A third possibility, cumulative damage, describes strength decreases which depend on the size as well as the number of stresses.

A) Cyclic damage is the simplest type to analyze for reliability calculations. A few examples will be given here to generalize those mentioned in section 3.

i) In case 3-a with independent stresses and known cycle-dependent strength y_n, it follows that the n-th sucess has probability.

$$p_{s_n} = \int_0^{y_n} f(x)dx \tag{5-1}$$

Thus, R(t) can be computed using (4-1) and

$$R_n = p_{s_1} R_{2,1} \cdots R_{n,n-1} \tag{5-2}$$

Even more cycle-dependence can be introduced here for situations in which the operating procedure is known

to produce different, say larger, stresses on later cycles. The cycle-dependent stress densities can be symbolized by f(x;n) and

$$R_{n,n-1} = \int_0^{y_n} f(x\ ;\ n)dx \qquad (5\text{-}3)$$

If the efficiency of a system decreases in proportion to the number of its cycles of operation, it will operate at a higher temperature during later cycles, and thus cause greater thermal stresses on adjacent parts. In such a case, the mean of f(x;n) would be an increasing function of n.

ii) Another interesting situation is version of 3-b with independent stresses, and a random fixed strength which has known cycle dependence. For example, the strength might be normally distributed with fixed variance σ^2 but cycle dependent mean

$$E\ [y_n] = a + b\ e^{-n} \qquad (5\text{-}4)$$

This mean value decreases from its initial value of (a+b) toward a final value of (a).

In this situation, the conditional densities in Section 3 discussion of this case must be replaced by

$$g_n(y;n) = P[\, y \leq y_n \leq y + dy \mid (n\text{-}1)\ \text{successes}] \ /\ dy \qquad (5\text{-}5)$$

in which the subscript indicates (n–1) previous successes, and the n in the argument indicates the cycle dependence. The corresponding change in the reliability derivation is to replace $g_1(y)$ by $g_1(y;n)$. This causes little additional difficulty in computing approximations to R(t), but prevents the notational simplicity of the final result since the order of summation and integration can no longer be reversed.

An alternate or additional cycle-dependence of stresses could be introduced into this Case 3-b situation. The required generalization is to replace f(x) by f(x;n). This requires changing the

$$[\ \int_o^y f(x)dx]^{n-1}$$

and related expressions to

$$[\ \int_o^y f(x;1)dx\ ]\ [\ \int_o^y f(x;2)dx]\ \ldots\ [\ \int_o^y f(x;n\text{-}1)dx] \qquad (5\text{-}6)$$

B) – Aging has effects which are very much like cycle dependence. However, the changes here depend on the stress-times t_n rather than on the stress-numbers n. Thus, the aging version of (5-3) is

$$R_{n,n-1}(t_n) = \int_o^{y(t_n)} f(x;t_n)dx \qquad (5\text{-}7)$$

In this case, the discrete reliability R_n corresponding to (5-2) is really a function of all previous stress times $t_1,\ t_2\ .\ .\ .\ t_n$:

$$R_n(t_1,t_2\ \ldots\ t_n = p_{s_1}(t_1)\ R_{2,1}(t_2)\ \ldots R_{n,n-1}(t_n) \qquad (5\text{-}8)$$

In like manner, the conditional densities like those defined in (5-5) take the even more complicated form.

$$g_n(y;t_1,\ t_2\ \ldots\ t_n) = P[\, y \leq y_n \leq y + dy \ [\text{successes at } t_1,\ t_2 \ldots t_n] \qquad (5\text{-}9)$$

6) Cumulative damage introduces much greater complexity, and the necessary calculations are not simply minor variations of the preceding discussion. Additional work must be done in order to find reliability functions for devices which suffer cumulative damage due to randomly occuring stresses. The points of view to be considered include the Palmgren-Miner rule which accounts for the size of previous stresses, but not their order of occurrence, as well as the work of Sweet and Kozen [7] which does not take into account the order of occurrence of previous stresses.

Of particular interest is the possibility that aging or cycle dependence models might provide simple but reasonably accurate approximations to more precise analysis of cumulative damage. In fact it is easy to develop a complete model for the specific case of cumulative sum damage described below.

If we assume that the cumulative damage weakens the strength of the part on each occurrence an amount proportional to the applied stress then we obtain a cumulative damage law based on the sum of the applied stresses.

$$y_2 = y_1 - cx_1$$

$$y_3 = y_2 - cx_2 = y_1 - cx_1 - cx_2$$

$$\cdots\cdots\cdots\cdots$$

$$y_n = y_1 - c\sum_{i=1}^{n} x_i \qquad (5\text{-}10)$$

Where c is the factor of proportionality. If we invoke the central limit law of probability and assume the distribution of stress does not change, we can state that the term $c\sum_{i=1}^{n} x_i$ will be normally distributed for n large and will have the moments

$$\text{mean} = n\ c_1\mu_x \qquad (5\text{-}11)$$

$$\text{variance} = n\ c_1^2\ \sigma_x^2 \qquad (5\text{-}12)$$

Using equations (5-10) – (5-12) we have essentially reduced the case of cumulative damage to a cycle dependent model.

Conclusions

The authors feel that this paper and the underlying work cited in the references provide techniques for modeling the reliability of non-electrical systems. It is recommended that these techniques be widely applied so that a bank of parameter data for such models can be amassed in the literature. This in some ways parallels the development of failure rate reliability models whose widespread application waited for the gathering of a failure rate data base.

It is also important that the various forms of

failure dependence be explored. To be more specific we discuss dependence in greater detail by focusing on the example of a space vehicle. Firstly, one may say that the stresses during boost are dependent since it is quite likely that one large vibration stress will be followed by another, yet if the vehicle survives until it is injected into orbit, the successive small stresses in orbit will be unrelated. Actually this is not a case of dependence but merely two modes of operation. Seperate reliability models should be written for the boost and orbit phases with different stress-strength parameters.

Now suppose that failures in the fuel feed system of our liquid fuel booster cause a pulsating fuel flow and large concommitant vibration stresses. This is a case of correlated failure modes and might be treated by a Markov model where the transition probabilities were adjusted to account for the dependence.

Lastly, let us suppose that while traveling through space on a deep space mission, our vehicle is suceptable to meteorite damage. Further assume that meteorite belts are of a small mass type, which do little or no damage, and of a large mass type. The vehicle can be disabled by one or more large mass hits. Now if we knew that there were no meteorite hits during the last minute of flight it would effect our computation of the reliability during the next minute. In fact, if meteorite belts took hours to traverse, the fact that we just entered a small mass belt a few minutes ago would increase the probability of survival (avoidance of a large mass belt) during the next minute. Similarly, if we just entered a large mass belt, the probably of survival during the next minute is decreased. Notice that in each case we are assuming some sort of direct or indirect measurement of the applied stress. (meteorite mass) In such a case a dependent model must be constructed and repair and replacement policies might be changed by the dependence. However, if we consider the case where we do not know and/or cannot infer the mass of the meteorite hits, we consider the stresses to be uncorrelated. Of course the effects of the large meteorite masses is not ignored. Either by studying past data or by analysis, the parameters of the uncorrelated model are chosen so that they include this effect. Thus, although the reliability calculated for a particular time interval may be less accurate, the expected values over a mission will still be correct.

It is the authors opinion that the application of the general reliability modeling techniques presented here is beyond the score of this paper. Here the purpose was to describe in detail theoretical aspects of the time and/or cycle dependence of the stress and strength distributions used to characterize failure modes in the Stress-Strength Interference Method.

A further paper (Part II) is planned for next year which will show how these modelling techniques are applied to practical engineering problems. This will be done through the use of examples drawn from previous design and analysis work of the authors.

Appendix - Derivations of (3.18) and (3.19)

This appendix completes the discussion of the case of random fixed strength and random independent stresses which was summarized in Section 3. The conditional strength density after (n-1) successes

$$g_n(y)dy = P[y < \underline{y} < y + dy \mid \underline{x}_i < \underline{y} \; ; \; i = 1, 2, \ldots n-1] \qquad \text{(A-1)}$$

can be written as the ratio of joint and marginal probabilities

$$g_n(y)\,dy = \frac{P[y < \underline{y} < y+dy \text{ and } \underline{x}_i < \underline{y} \; ; \; i = 1, 2, \ldots n-1]}{R_{n-1}} \qquad \text{(A-2)}$$

where the reliability R_{n-1} is the probability of (n-1) consequtive successes.

The numerator of (A-2) can be expanded in terms of additional conditional probabilities to the form

$$P[y < \underline{y} < y + dy]\; P[\underline{x}_{n-1} < \underline{y} \mid y < \underline{y} < y + dy \text{ and } (n-2) \text{ successes}] \times$$

$$\times\; P[\underline{x}_{n-2} < \underline{y} \mid y < \underline{y} < y + dy \text{ and } (n-3) \text{ successes}] \times$$

$$\vdots$$

$$P[\underline{x}_1 < \underline{y} \mid y < \underline{y} < y + dy \qquad \text{(A-3)}$$

The assumed independence of stresses makes all of those conditional probabilities equal to

$$\int_0^y f(x)dx \qquad \text{(A-4)}$$

Substitution of (A-4) and (A-3) into (A-2) produces the conditional strength density shown in (3.18). Finally, (3.19) follows directly from substituting (3.18) into (3.16). Similar derivations appear in Ref. 1,2,8.

REFERENCES

1. A.M. Freudenthal, et al, "The Analysis of Structural Safety," J of the Structural Div., ASCE, Vol. 92, No. 572, Feb. 1966, p. 267-325.

2. G. Reethof, et al, "A model for Time Varying and Interfering Stress-Strength Probability Density Distributions with Consideration for Failure Incidence and Property Degradation," SAE-ASME-AIAA Aerospace Reliability and Maintainability Symposium, June 1964, pp. 566-575.

3. M. Shooman, "Probabilistic Reliability: An Engineering Approach", McGraw-Hill, 1968.

4. R. Schatz, "Time/Cycle Reliability Modeling Techniques utilizing the Stress-Strength Interference Method, Westinghouse Astronuclear Laboratory memo No. 54071.

5. R. L. Disney, C. Lipson, and N. J. Sheth, "The Determination of the Probability of Failure by Stress/Strength Interference Theory, IEEE Symposium on Reliability, January 1968.

6. H. Cramer and M. R. Leadbetter, Stationary and Related Stoclastic Processes, Wiley, N.Y., 1967 .

7. A. L. Sweet and F. Kozin - "Investigation of a Random Cumulative Damage Theory", Journel of Materials, Vol. 3, No. 4, 1968.

8. G. Wagner, "Procedures for Structural Reliability Analysis", Westinghouse Astronuclear Laboratory Memo. No. 52296.

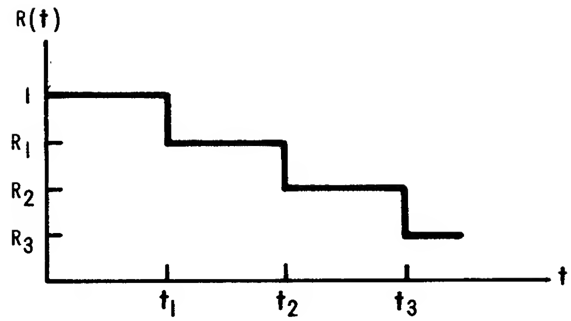


Figure 3.1. Reliability Function for Repeated Stresses

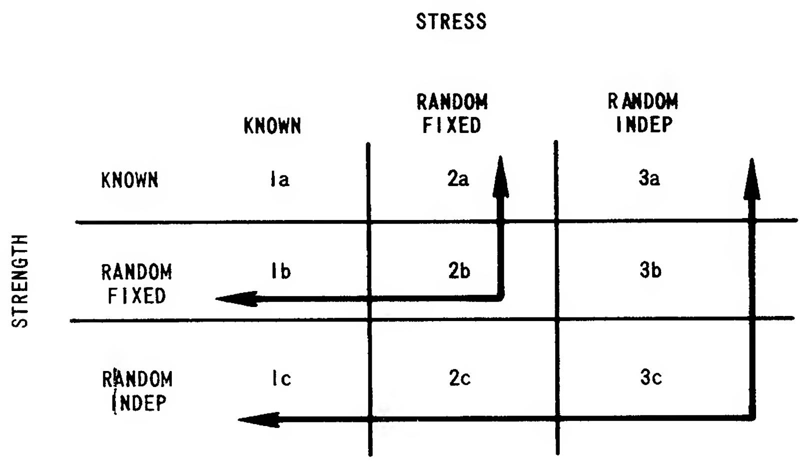


Figure 3.2. Diagram of Cases

5208-1

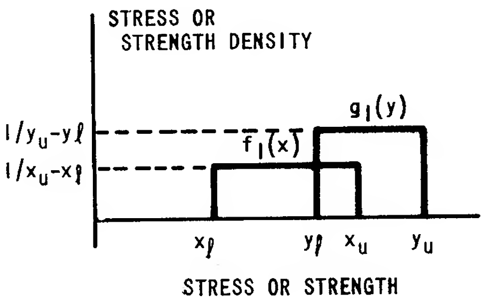


Figure 3.3. Rectangular Stress and Strength Densities

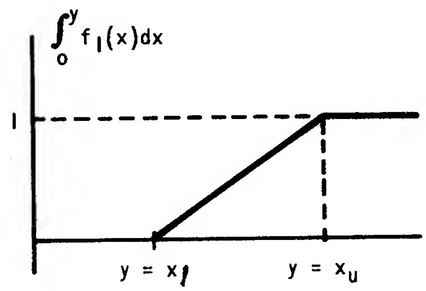


Figure 3.4. Computation of $\int_0^y f_1(x)dx$

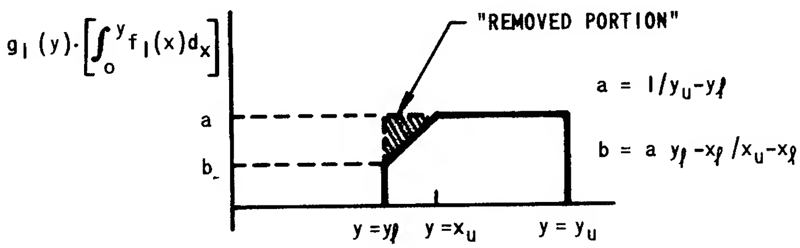


Figure 3.5. Computation of Numerator of Eq 3.21

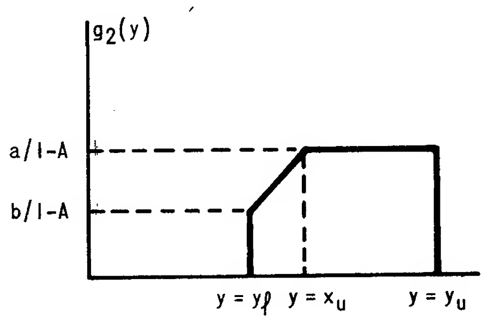


Figure 3.6. Computation of $g_2(y)$

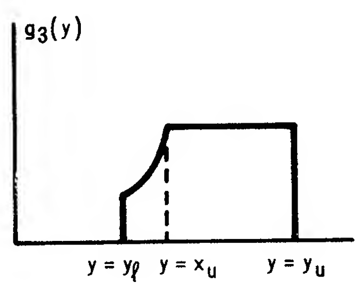


Figure 3.7. Shape of $g_3(y)$

5208-2

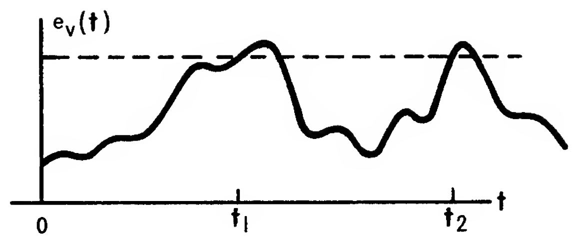


Figure 4.1. Gaussian Threshold Crossings

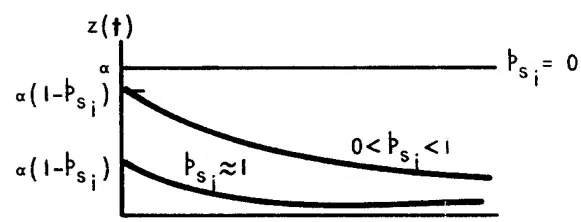


Figure 4.2. Hazards for Case 3-a

5208-3

ANALYSIS OF EARLY FAILURES IN UNEQUAL SIZE SAMPLES

by

R. A. Heller and A. S. Heller
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Summary

Early failures in unequal size fleets are used to estimate the population distribution function using a combination of analytical and graphical methods. The analysis is based on order statistics and extreme value theory.

List of Symbols

g_i	number of identical size samples
i	index
k	rank order number
m	number of sample sizes
n, n_i, $\bar{n}$	sample size, for ith sample, average sample size
p_i, $p(n_i)$	frequency, density of sample sizes
x	general variate
E	expected value
$F(x)$	cumulative probability function
M	total number of samples
N	fatigue life, number of cycles
$R(x)$, $R(N)$, $R(\rho)$	reliability function of variate x, or N, relation between R and ρ
ρ_{1n_i}, ρ_{2n_i}	first and second failure reliability functions for variable fleet n_i
$\rho_1(N)$, $\rho_2(N)$	first and second failure reliability functions for fatigue life N

Introduction

Sampling tests for the purposes of quality control, for test acceleration, or for the detection of weak members of a population may benefit greatly from a particular use of the statistics of extremes. It has been shown by the authors in a previous paper[1] that it is expedient to perform tests on groups of specimens taken from a larger population until the weakest (first failure) in each group, or fleet, fails. The surviving specimens are then not tested any further and population parameters are estimated based on a knowledge of the distribution of first failures and the size of the samples tested.

Because testing is discontinued after the first failure has occurred in each sample an obvious acceleration in testing time is realized. When sampling of this nature is performed it is expedient to use groups of identical size for which case the techniques have been established in reference 1. The method has recently been utilized in fatigue tests of 17 large panels each containing 32 rivet holes[2]. The holes and surrounding material were considered to be members of 17 samples of size 32. Tests were discontinued after the appearance of the first crack in each panel and these first failures were used to estimate the fatigue life distribution for all holes.

It is, however, of great practical interest to extend the analysis of early failures to samples of unequal size. An aircraft manufacturer for instance may sell different numbers of airplanes to the various airlines. As these unequal size fleets undergo service, the first failures in each fleet may be utilized to estimate life parameters for the surviving population.

A non-parametric method is presented here for the graphical estimation of the population distribution function using the first and second failures in unequal size samples.

Reliability Functions of Early Failures

From the theory of extreme value statistics[3] it is well known that in a sample of size "n", drawn from a given population with reliability function, R(x), the smallest value of x will have a reliability function

$$\rho(x)_{1n} = R(x)^n \tag{1}$$

whether the sample consists of one or more sets of drawings, all of size "n". Generally, the "k"-th smallest rank in the sample will have a reliability function, ρ_{kn}, equal to the sum of the last k terms of the binomial expansion[1] denoted here as

$$\rho_{kn} = \{[R(x) + F(x)]^n\}_k$$
$$= \sum_{i=n-k+1}^{n} \binom{n}{i} R(x)^i F(x)^{(n-i)} \tag{2}$$

with F(x) the failure function of the parent population. Equations 1 and 2 are nonparametric and relate the reliabilities of the sample and that of the parent population without requiring a prior knowledge of the parent distribution.

In this paper the more general case, that of unequal size samples, is treated. This altered situation, however, adds difficulties to the derivation and solution of the equations.

Several underlying assumptions are noted below:

(a) The sample size "n" is <u>not</u> a constant, but a random variable.
(b) The distribution function of the random variable "n" is necessary for a solution of the reliability equations.
(c) The reliability functions of the "k"-th smallest values depend on the distribution of sample sizes, and hence conditional probabilities must be used as follows.

A set of M samples of various sizes n_1, n_2 ... n_m are chosen. It is assumed that $n_1 \neq n_2 \neq n_i$; that is, <u>none</u> of the samples are of the same size.

In each ith sample there are n_i ranked values with a reliability function dependent on the rank of the variate and the size of the sample, n_i Using the format of Equation 1 above, the reliability function of the smallest value for a sample of size n_i is defined as

$$\rho_{1n_i} = R(x)^{n_i}\Big|\, n = n_i \text{ for } i = 1, 2 \ldots m \tag{3}$$

given that the size of the sample n equals n_i. Choosing at random, among the M samples of sizes n_1, n_2 ... n_m, it is possible to develop the reliability function of the "k"-th smallest value for a particular distri-

bution of sample sizes characterized by the density function $p(n_i)$. When the density function, $p(n_i)$, is known, the mean of the sample sizes is

$$E(n) = \sum_{i=1}^{n} n_i\, p(n_i)$$

$$= \sum_{i=1}^{n} n_i [p(n_i) \big| n = n_i] \qquad (4)$$

Thus the reliability function of the smallest value for sample sizes with a density function $p(n_i)$ can be unconditionally defined as

$$\rho_{1p(n_i)} = \sum_{i=1}^{m} p(n_i) [R(x)^{n_i} \big| n = n_i] \qquad (5)$$

If $p(n_i)$ is a known function, Equation 5 can be readily evaluated even if the R(x) are not assumed a priori, keeping the relationship between the sample and the parent population nonparametric. In practice, however, instead of the density function, $p(n_i)$, the frequency of occurrence of each sample size is available and may be defined as

$$p_i = g_i/M, \; i = 1, 2 \ldots m \qquad (6)$$

where g_i is the number of times a sample of size n_i occurs, M is the total number of samples, and m is the number of different sample sizes.

Of course

$$\sum_{i=1}^{m} p_i = 1 \text{ and } \sum_{i=1}^{m} g_i = M \qquad (7)$$

Hence a weighted average sample size, $\bar{n}$, may be evaluated based on Equation 4 as

$$\bar{n} = \sum_{i=1}^{m} p_i n_i \qquad (8)$$

Thus based on Equation 5 the unconditional reliability function of the smallest values in the average sample size $\bar{n}$ is

$$\rho_{1\bar{n}} = \sum_{i=1}^{m} p_i R(x)^{n_i} \big| n = n_i \qquad (9)$$

It should be noted that the same $\bar{n}$ may be obtained with various combinations of p_i and n_i and hence it is incorrect to assume that $\rho_{1\bar{n}}$ could be computed from Equation 1 by substituting $\bar{n}$ for n. In other words

$$\sum_{i=1}^{m} p_i\, R(x)^{n_i} \neq R(x)^{\bar{n}} \qquad (10)$$

However, as will be shown later, the left and right sides of Equation 10 will be approximately equal for high reliabilities.

Having established Equation 9 the reliability function of the kth smallest value in a sample of size $\bar{n}$ readily follows

$$\rho_{k\bar{n}} = \sum_{i=1}^{m} p_i \{[R(x) + F(x)]_k^{n_i} \big| n = n_i\} \qquad (11)$$

with the subscript k referring to the last k terms of the expansion of $[R(x) + F(x)]^{n_i}$.

Thus for k = 2, Equation 11 can be rewritten as

$$\rho_{2\bar{n}} = \sum_{i=1}^{m} p_i \{[R(x)^{n_i} + \binom{n_i}{1} R(x)^{(n_i-1)} F(x)] \big| n = n_i\} \qquad (12)$$

For k = 3, the last three terms of Equation 11 yield $\rho_{3\bar{n}}$. It should be recognized that Equation 11 is valid only for k values not larger than the smallest of the sample sizes, n_i. If $g_i = 1$ for all i that is each sample size occurs once, p_i is reduced to $\frac{1}{M}$ and M = m; for k = 1

$$\rho_{1,\bar{n}} = \sum_{i=1}^{m=M} \frac{1}{M} R(x)^{n_i} \big| n = n_i, \text{ and} \qquad (13)$$

for k = 2

$$\rho_{2\bar{n}} = \sum_{i=1}^{m=M} \frac{1}{M}[R(x)^{n_i} + \binom{n_i}{1} R(x)^{(n_i-1)} F(x) \big| n = n_i] \qquad (14)$$

which are special cases.

Additionally, should all the samples be of the same size, that is $n_i = n$, the summation of Eq. (9) will contain only one term ($i = 1$, $p_i = 1$) and Eq. (9) will reduce to Eq. (1), which is for equal size samples.

Prediction of the Population Distribution

To illustrate the use and validity of the foregoing, the same 95 fatigue test results[4] examined in reference 1 with equal sample sizes will be reevaluated here using unequal fleets.

The 95 rotating bending fatigue lifetimes, N, for 7075-T6 aluminum at a stress of 37,300 psi are presented in Table I. They were ranked in increasing order and were plotted on extreme value probability paper to provide the reliability-life distribution, R(N), of the "parent" population. The curve is shown in Figure 1. The mean plotting position, $1 - \frac{k}{n+1}$, was used for the estimate of the reliability on all figures.

The same data were then randomly placed into various "fleets". The smallest values in these fleets were subsequently ranked in increasing order and were again plotted on extreme value probability paper as first failure reliability-life curves. The same procedure was employed using the second smallest values to obtain second failure curves as shown on Figure 1.

These first and second failure distributions were then used to predict the distribution function of the "parent" population which was eventually compared to the original distribution function for the complete set of 95 test results.

Eqs. (9) and (12) were evaluated for the particular fleet distributions shown on Figs. 1, 2 and 3 with the aid of an electronic computer. Values of R ranging from .9999 to .1000 were chosen and the corresponding values of ρ_1 and ρ_2 were calculated and plotted as $R(\rho_1)$ and $R(\rho_2)$ curves on the left hand sides of Figs. 1, 2, and 3. These relationships, being nonparametric, could have been plotted on any graph paper but, in order to facilitate the prediction of the "parent" dis-

Table I. Constant Stress Amplitude Tests on 7075-T6 Aluminum at ±37,300 psi

N No. of Cycles to Failure in Hundreds									
No. k	N	No. k	N	No. k	N	No. k	N	No. k	N
1	5136	21	10104	41	12403	61	13967	81	18087
2	6076	22	10154	42	12409	62	14055	82	18293
3	6249	23	10451	43	12434	63	14283	83	18643
4	7425	24	10512	44	12449	64	14432	84	18898
5	7680	25	10667	45	12482	65	14558	85	19121
6	7993	26	10785	46	12488	66	15035	86	19129
7	8249	27	10834	47	12531	67	15036	87	20087
8	8816	28	11066	48	12629	68	15091	88	20281
9	8853	29	11255	49	12723	69	15135	89	20518
10	8878	30	11319	50	12823	70	15135	90	20798
11	9114	31	11490	51	12920	71	15165	91	21118
12	9268	32	11597	52	12921	72	15416	92	21231
13	9273	33	11600	53	13010	73	15485	93	21794
14	9297	34	11623	54	13017	74	16117	94	23318
15	9315	35	11799	55	13171	75	16372	95	27319
16	9481	36	11864	56	13185	76	16452		
17	9484	37	11942	57	13234	77	16872		
18	9662	38	11969	58	13309	78	17461		
19	9666	39	12211	59	13352	79	17827		
20	10058	40	12323	60	13824	80	18025		

tribution, the same extreme value probability scale was used for both R and ρ as for the reliability-life curves.

The parent population reliability curve was reconstructed from the first and second failure curves as follows: a horizontal line was drawn from a point on the $\rho_1(N)$ curve of Fig. 1 towards the left until it intersected the 45° line, from there a vertical line intersected the $R(\rho_1)$ curve at the required reliability value R; a horizontal line was then drawn towards the right back to the original value of N to locate a point on the R(N) line. This procedure was repeated for $\rho_2(N)$.

In this manner all first and second failure points were moved up to new positions. A line drawn through these points approximates the original parent population reliability curve very well.

In Figure 1, the 95 data were subdivided into one fleet of 50, one of 20, two of 10 and 1 of 5 members. Hence the five fleets provided only 5 first failure and 5 second failure points to approximate the parent distribution function.

For Fig. 2, 90 of the 95 fatigue lives were placed into 9 fleets: four with 5 members, three with 10, and two with 20. Having a larger number of first and second failures available the approximation to the parent population reliability curve is much better than in the case shown in Fig. 1.

The average fleet size for Fig. 2 based on Eq. 8 is $\bar{n} = 10$. For the sake of comparison the $R(\rho_1)$ and $R(\rho_2)$ curves for a constant sample size of ten were computed [Eqs. (1) and (2)] and are also shown on the left side of Fig. 2. It is seen that the use of the average fleet size would lead to errors at low levels of reliability, as indicated by Eq. (10), while for higher values of R Eq. (10) can be used as an approximate equality.

In Fig. 3, two fleets consisting of single members were chosen in addition to four fleets of 5, three samples of 10, and two of 20. Because the two fleets with one member each are exhausted after the first failure, no second failure curve can be plotted.

The procedure, however, works very well even in this special case as indicated by the good fit of the estimated reliability points around the "parent" population curve.

Recurrence Relations for Equal Size Samples

It is useful to note that successive failure probabilities are related to each other through recurrence relations. As a result the calculation of the reliability of second failures in samples of constant size n_i can be simplified by determining instead, the reliability of first failures in a sample of reduced size, n_i-1.

Generally for the kth rank in a sample of size n_i [5]

$$\rho_{k(n_i-1)} = \frac{n_i-k}{n_i}\rho_{kn_i} + \frac{k}{n_i}\rho_{(k+1)n_i} \qquad k = 1, 2, \ldots, n_i-1 \tag{15}$$

To prove the validity of Eq. (15) for k = 1 for instance, Eqs. (1) and (2) will be used:

$$\rho_{1n_i} = R(x)^{n_i} \tag{16}$$

and

$$\rho_{2n_i} = R(x)^{n_i} + n_i R(x)^{(n_i-1)}[1 - R(x)] \tag{17}$$

Substituting Eqs. (16) and (17) into Eq. (15) yields

$$\rho_{1(n_i-1)} = R(x)^{(n_i-1)} \tag{18}$$

which is the first failure in a sample of size $n_i - 1$ based on Eq. (1).

Eq. (15) is valid for a single sample of size n_i or for several samples of equal size n_i. Using Eq. (15) the reliability of second failures ρ_{2n_i} may be expressed as

$$\rho_{2n_i} = n_i \rho_{1(n_i-1)} - (n - 1)\rho_{1n_i} \tag{19}$$

Because reliabilities with unequal sample sizes involve the weighting term, p_i, no recurrence relation analogous to Eq. (19) has been found for variable n_i. However, the average sample size $\bar{n}$ of Eq. (8) might be used in Eq. (19) for an approximation of $\rho_{2\bar{n}}$.

Should such an approximation prove to be close, it would be useful in the estimation of second failure reliabilities for cases similar to the one presented in Fig. 3. In that example, some of the samples become exhausted after the first failure and hence an exact calculation of $R(\rho_2)$ is not possible.

An equation analogous to Eq. (15) with k = 1 can be written for variable fleet size

$$\rho_{1(\bar{n}-1)} = \sum_{i=1}^{m} \left\{ p_i \left(\frac{n_i - 1}{n_i}\right) R(x)^{n_i} + \frac{1}{n_i}\left[p_i R(x)^{n_i}\right. \right.$$

$$\left.\left. + \, p_i n_i R(x)^{(n_i-1)} \left[1 - R(x)\right]\right]\right\} \tag{20}$$

but the individual terms $\rho_{1\bar{n}}$ and $\rho_{2\bar{n}}$ cannot be separated. Eq. (20), after simplification, does reduce to the correct form

$$\rho_{1(\bar{n}-1)} = \sum_{i=1}^{m} p_i R(x)^{(n_i-1)} \tag{21}$$

which is similar to Eq. (5).

Conclusions

The technique of testing to first and second failures for quality control, for material properties,[1] and for structural[2,6] and mechanical integrity has become a useful tool recently. It is now possible to perform this type of testing and analysis even when widely different sample sizes are used.

The work reported here has been supported by the U.S. Air Force Materials Laboratory under contract no. AF33615-72-C-2111.

References

[1] Heller, R. A. and Heller, A.S., "The Relationship of Earliest Failures to Fleet Size and Parent Population," Annals of Reliability and Maintainability, V. 5, pg. 722 (1966).

[2] Eggwertz, S., "Investigation of Fatigue Life and Residual Strength of Wing Panel for Reliability Purposes," ASTM, STP 511, pg. 75 (1972).

[3] Gumbel, E. J., "Statistics of Extremes," Columbia University Press, New York, pg. 78 (1958).

[4] Freudenthal, A. M., Heller, R. A. and O'Leary, P. J., "Cumulative Fatigue Damage of Aircraft Structural Materials," Air Force Systems Command WADC TN 55-273, pt. I, (1955).

[5] David, A., Order Statistics, John Wiley and Sons, New York, pg. 37 (1970).

[6] Freudenthal, A. M., "Expected Time to First Failure," Air Force Systems Command AF ML 66-37 (1966).

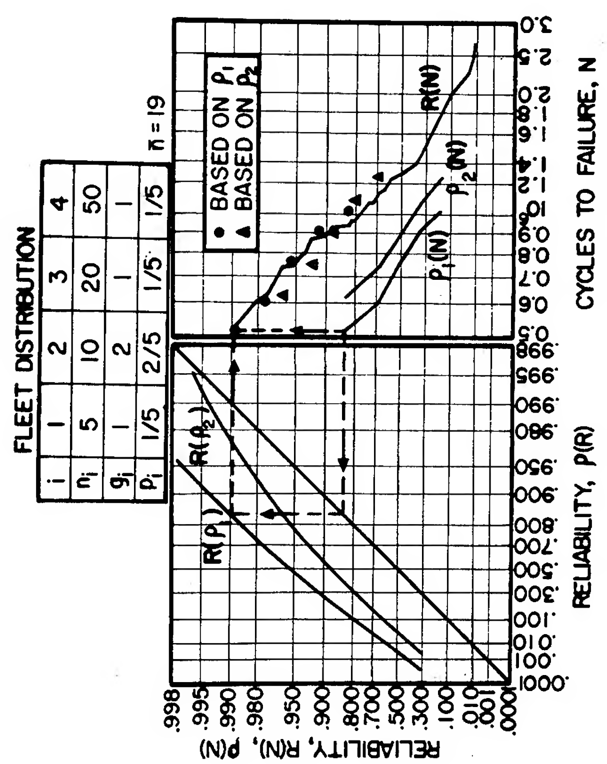


Fig. 1. Early failure and population distributions; $\bar{n}$ = 19.

FLEET DISTRIBUTION

i	1	2	3
n_i	5	10	20
g_i	4	3	2
p_i	4/9	3/9	2/9

$\bar{n} = 10$

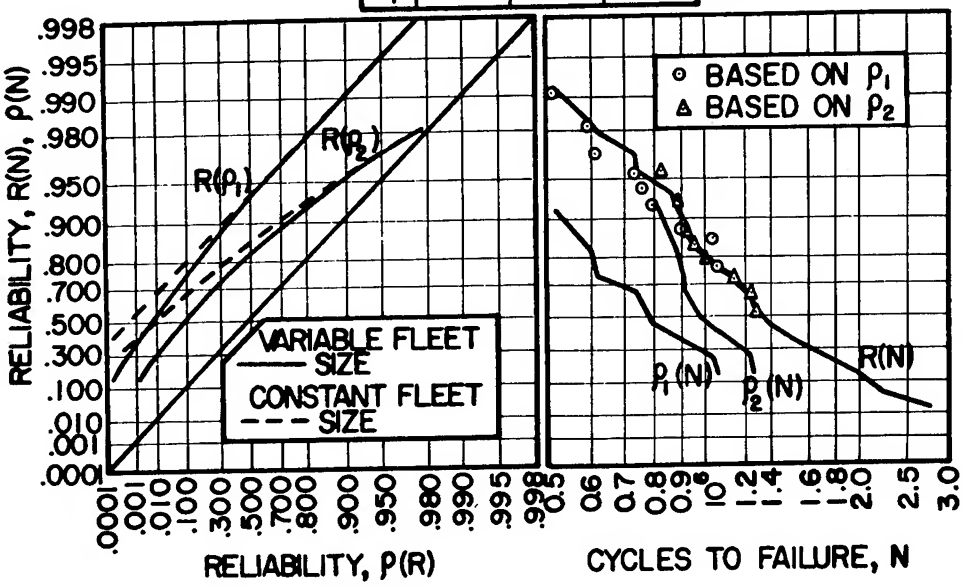


Fig. 2. Early failure and population distributions; $\bar{n}$ = 10.

FLEET DISTRIBUTION

i	1	2	3	4
n_i	1	5	10	20
g_i	2	4	3	2
p_i	2/11	4/11	3/11	2/11

$\bar{n} = 8.36$

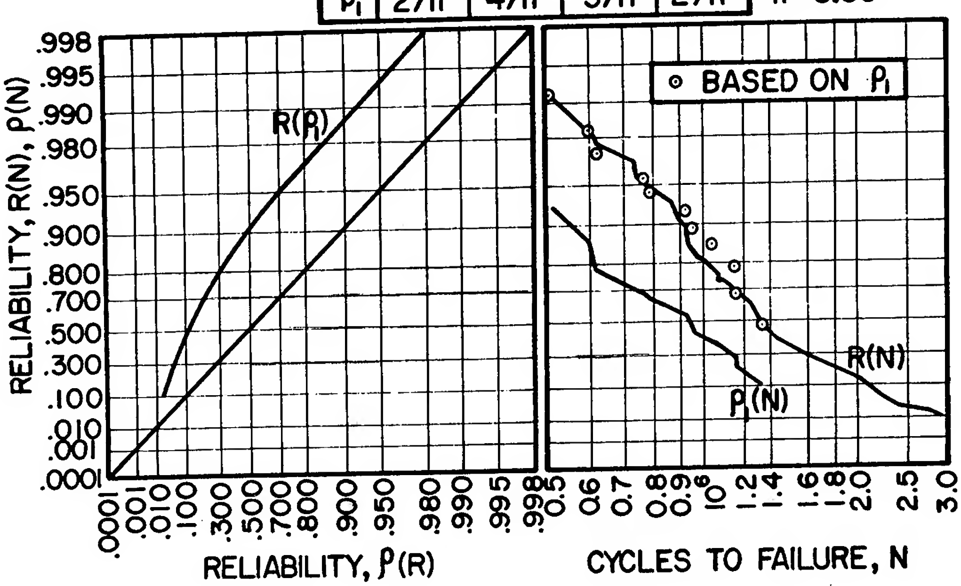


Fig. 3. Early failure and population distributions; $\bar{n}$ = 8.36.

RELIABILITY OF $GaAs_{1-x}P_x$ LIGHT EMITTING DIODES

Shin-Ichiro Koyama, Toshihiko Kohisa and Shigeru Suzuki

Semiconductor Division, Hitachi Ltd.,
Kodaira-shi, Tokyo, Japan

Abstract

The life and environmental tests on Zn diffused $GaAs_{1-x}P_x$ light emitting diodes were carried out and the results was that MTBF exceeded 10^4 hours. It is observed that degradation of the light output increases with current density and temperature under forward bias operation. Life time of light emitting diodes decreases in proportion to $J_F^{-1.0}$ and the activation energy of life time is evaluated to be nearly 0.4eV. It is noticed that life time decreases rapidly with increasing peak current density in pulse operation even if averaged current density is maintained constant.

1. Introduction

The reliability of light emitting diodes has not yet been appraised sufficiently in the market in comparison with other semiconductor devices, since they are comparatively new devices. The various life and environmental tests have shown them to have as high reliability as transistors and integrated circuits apart from degradation of the light output.

It is well known that the light output decreases in a somewhat exponential manner with increasing temperature and current density under forward bias operation.

Several models of degradation in III - V light emitting diodes have been proposed and studied.[1] However, degradation mechanisms have been investigated only at rather high current density or under extreme condition. The useful analysis of degradation under practical conditions has not been reported in any detail.

In this report, we investigated reliability of Zn diffused $GaAs_{1-x}P_x$ light emitting diodes, especially the relation between degradation and current density or temperature.

2. Experiments

$GaAs_{1-x}P_x$ light emitting diodes studied in this work were Zn diffused and all were phospho-silicate glass passivated. The junction was a few micron in depth. They were encapsulated with transparent dome-shaped plastic as shown in Fig. 1 and some diodes were mounted on TO-18 headers without any encapsulation.

Reliability tests such as life and environmental tests were conducted on plastic encapsulated diodes. Degradation of the light output was investigated under the following forward bias operation;

(1) d. c. operation, at room temperature stressed at 55.6, 111 and 167A/cm^2, and 80°C, 100°C and 120°C stressed at 11A/cm^2 on plastic encapsulated diodes.
(2) pulse operation at room temperature with peak current density ranging from 148 to 1480A/cm^2 under the next conditions,
   (A) pulse width of 50μs and duty cycle of 0.05 on diodes without any encapsulation.
   (B) pulse width of 25μs with maintained average current density constant (74A/cm^2) on plastic encapsulated diodes.

The light output was determined by inserting the diode into the open side of a hollow cube composed of silicon solar cells.

3. Results and discussions

(1) The failure in the light output and the effect of the thermal expansion of plastic were examined under various accelerated life and environmental conditions. Quite satisfactory results were obtained as shown in Table 1. MTBF was estimated to exceed 10^4 hours at the 90% confidence level under these severe conditions.

(2) It is generally accepted that the time when the light output decreases to one half of the initial valve, is called the life time.

Degradation of the light output at room temperature is shown in Fig. 2. Life time is shown in Fig. 3, which was calculated assuming that the light output would decrease in a exponential manner except for the initial stage of degradation.

The results at high ambient temperature stressed at 11A/cm^2 are shown in Fig. 4. The activation energy of life time was estimated to be nearly 0.4eV. As a result, life time at room temperature is expected to reach at nearly 10^5 hours.

It was further concluded that at fixed junction temperature, current density was the dominant factor in degradation rather than ambient temperature in the region of practical junction temperature.

It was difficult in d. c. operation to obtain the relation between life time and current density due to the increase of the junction temperature. Therefore, pulse operation was set up on diodes without any encapsulation with pulse width of 50μs and duty cycle of 0. 05, to minimize the effect of heating. The results are shown in Fig. 5. The relation between life time ($t_{1/2}$) and current density (J_F) is given by,

$$t_{1/2} = \alpha \cdot J_F^{-1.0}$$

where α is a constant. Life time at peak current density larger than 10^3 A/cm^2 and that at the smallest did not apply to the above formula. The former was considered to be due to heating. As the degradation rate was extremely small, the latter was considered to have been shortened by the accidential error in approximating degradation according to the exponential decrease.

Next, as shown in Fig. 6, life time decreased with increasing peak current density in comparison with d. c. operation, even if average current density was kept constant (74A/cm^2). The first reason for these results is the effect of heating, however, the increase of the junction temperature was considered to be rather gradual as averaged current density was maintained constant. The rapid decrease of life time could not be explained only by heating. Therefore, it was noticed that peak current density played an effective role in degradation in addition to the effect of heating, that is, more electrons being injected into the P region than with d. c. operation was a possible cause for a more rapid degradation rate even if averaged current density was maintained constant.

(3) Degradation of light output is attributed to one of the external quantum efficiency. The external quantum efficiency η_{exter} is given by,

$$\eta_{exter} = \eta_0 \cdot \gamma \cdot \eta_{bulk}$$

where η_0 includes relative luminosity coefficient and internal absorption probability and γ and η_{bulk} are injection efficiency of electrons into P region and bulk luminescence efficiency.

Current-voltage characteristics and light output-voltage characteristics before and after aging stressed at 250°C and 15 A/cm^2, are shown in Fig. 7. Current-voltage characteristics after aging were shifted to the left in parallel with the initial one. Light output-voltage characteristics were shifted to the right in a similar way. This increase of current at fixed voltage have been attributed to the one of recombination current in the space charge region or excess current.[1] This change in current-voltage characteristics can account for degradation of injection efficiency at fixed current, namely, degradation of the external quantum efficiency.

4. Conclusions

The various life and environmental tests verified the high reliability of plastic encapsulated light emitting diodes. We determined the relation between life time and current density, and junction temperature. Life time decreased in proportion to $J_F^{-1.0}$ and the activation energy of life time was obtained to be nearly 0. 4eV. Life time at room temperature with current density of 11A/cm^2 was estimated to be nearly 10^5 hours. It was found that even if averaged current density was kept constant in pulse operation, life time decreased rapidly with increasing peak current density.

Acknowledgment

We express our sincere appreciations to Eiichi Adachi, Hitachi Central Lab. for many valuable suggestions.

Reference

1) Discussions of degradation models are given in the review by, L. R. Weisberg, IEEE Reliability Physics Symposium, April 14 (1970). and general discussions are given by, for example, A. A. Bergh and P. J. Dean, Proc. IEEE, 60, 156 (1972)

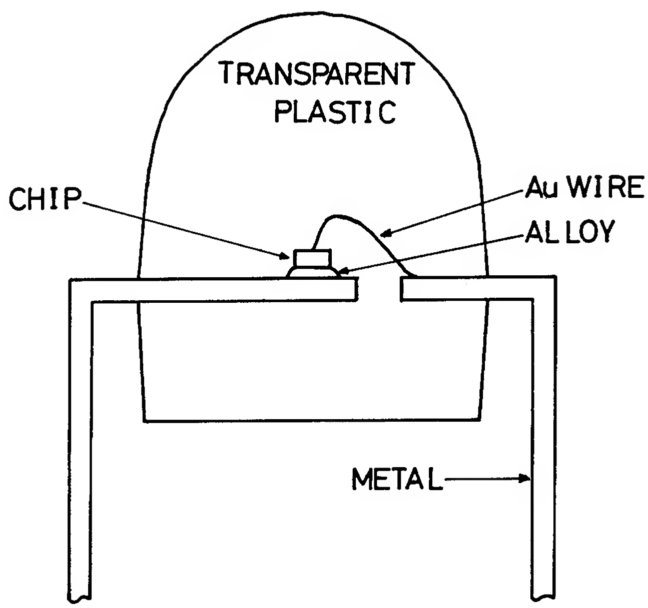


Fig. 1 The configuration of plastic encapsulated light emitting diodes.

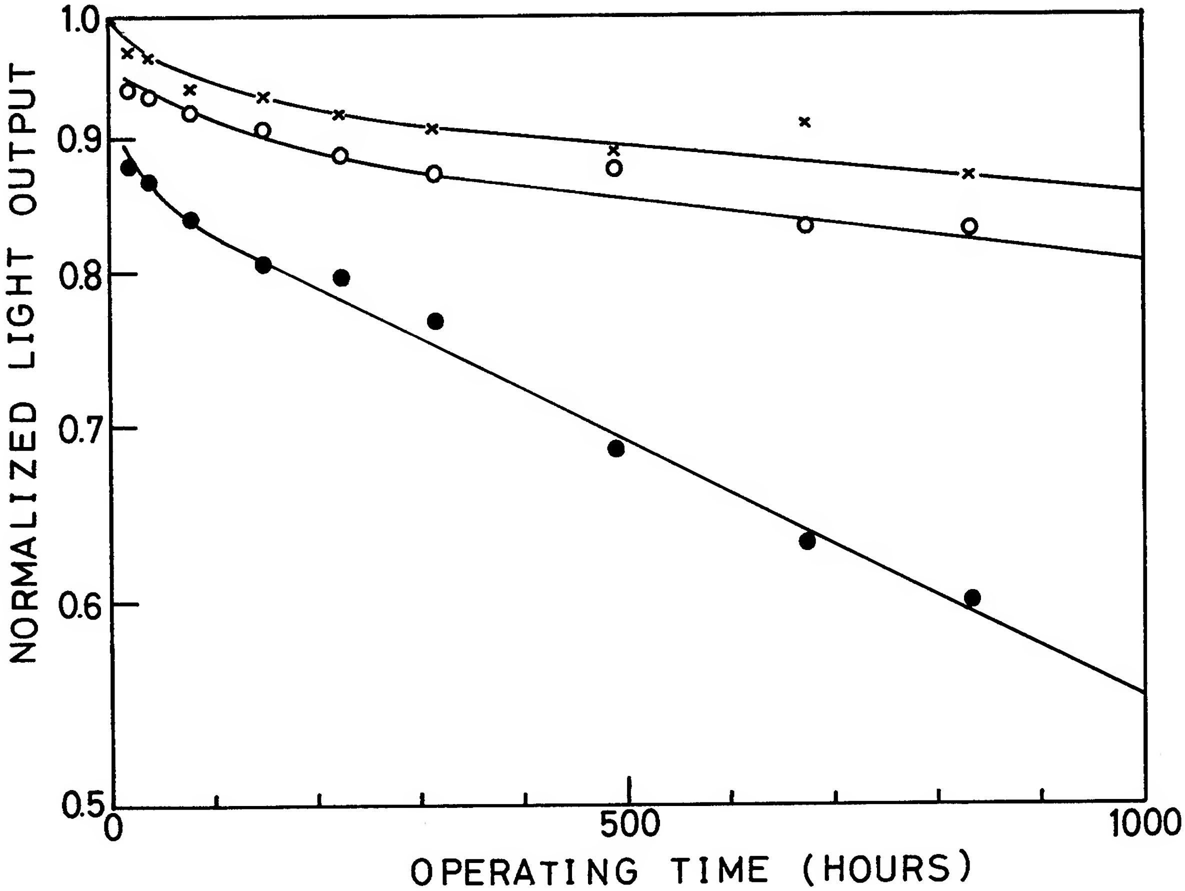


Fig. 2 The time change in the normalized light output at room temperature stressed at 55.6 (crosses), 111 (open circles) and $167 A/cm^2$ (closed circles). Each data point represents the mean of a group of ten diodes.

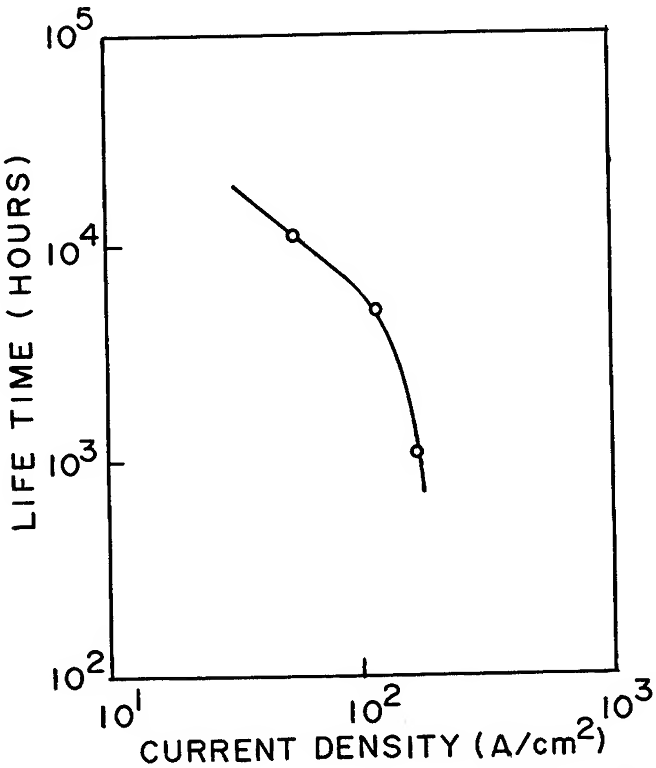


Fig. 3 Life time versus current density at room temperature.

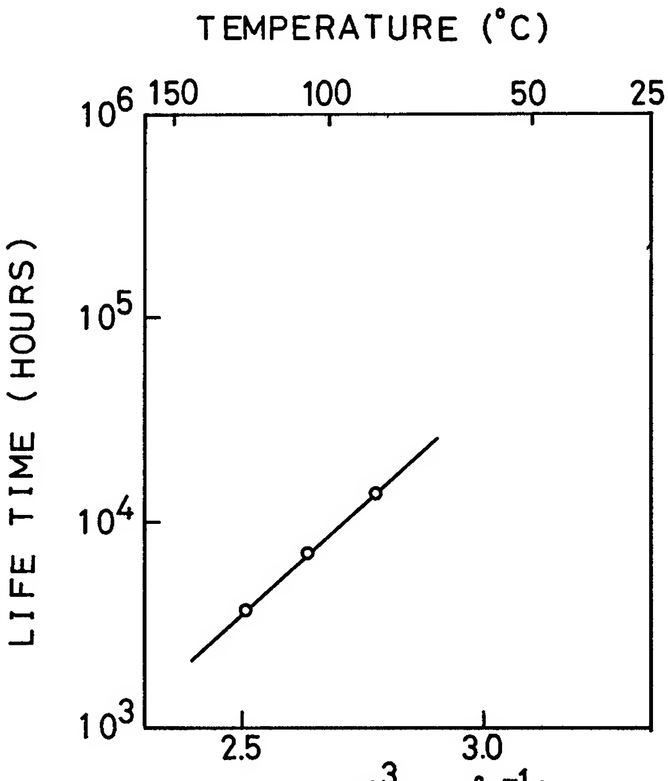


Fig. 4 Life time versus $10^3/T$°K at $11 A/cm^2$.

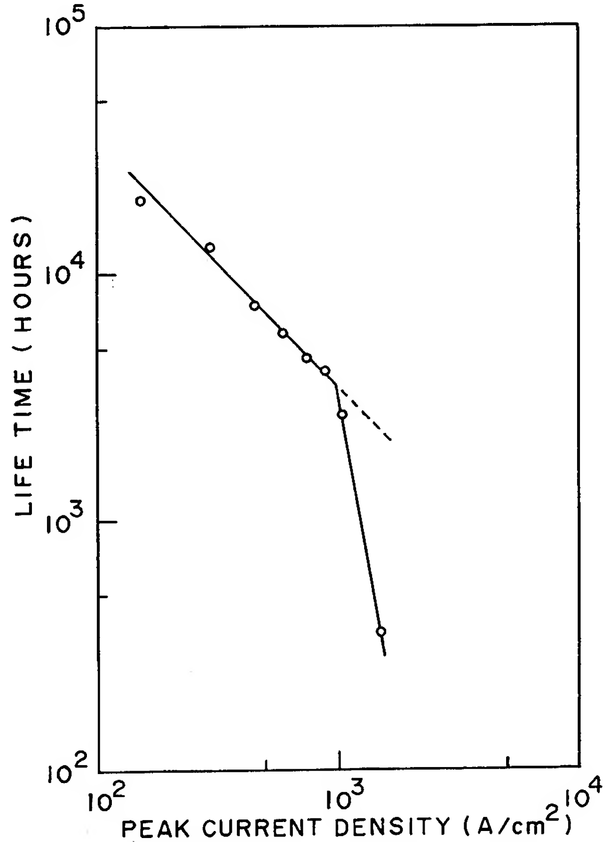


Fig. 5 Life time versus peak current density at room temperature; pulse width and duty cycle are 50us and 0.05, respectively.

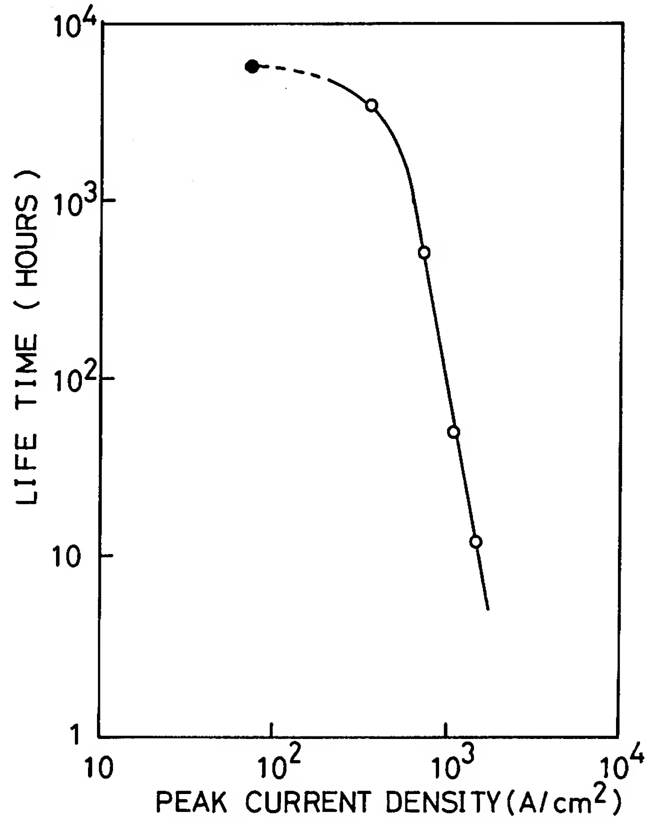


Fig. 6 Life time versus peak current density at room temperature. Averaged current density is maintained $74A/cm^2$ with pulse width of 25us. Life time at $74A/cm^2$ (d. c.) is shown (closed circle) for comparison.

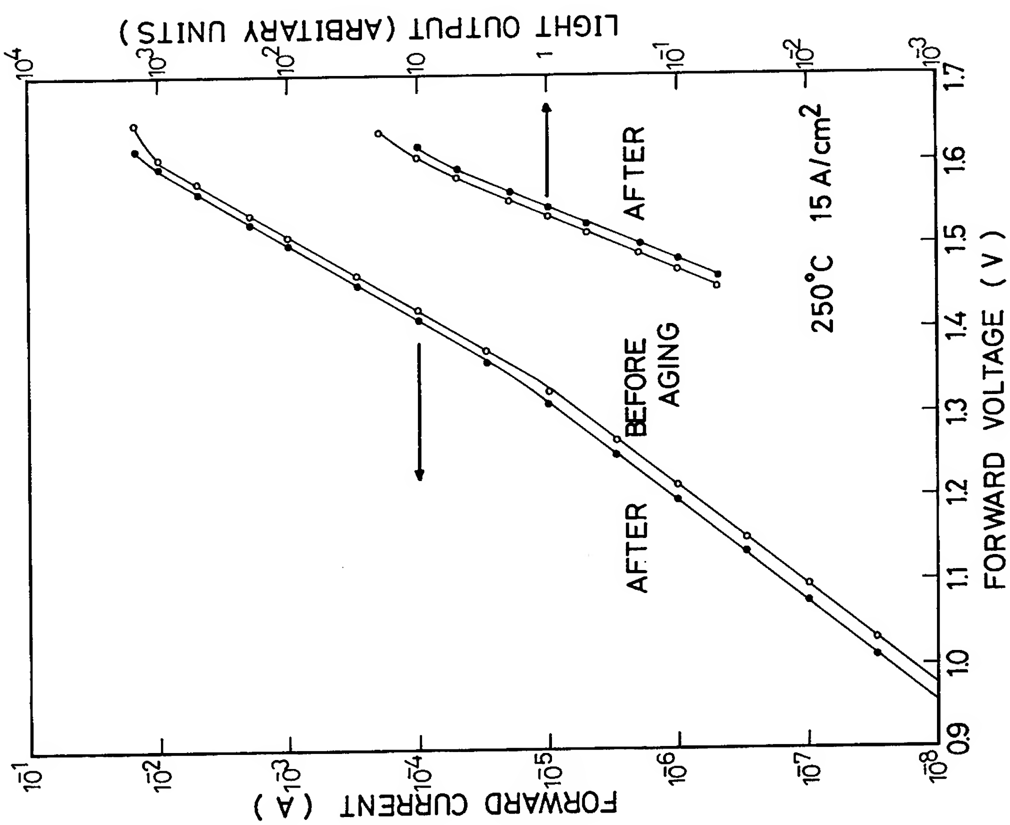


Fig. 7 Forward current versus voltage (in the left-hand side) and light output versus voltage (in the right-hand side), before (open circles) and after degradation (closed circles) at 250° C and $15A/cm^2$.

Table. 1 The results of life and environmental tests.

(1) The results of life tests

items	test conditions	sample quantity	total component-hour	failure	failure mode	failure rate (C. L. 90%)
forward bias	Pc=90mW	147	226, 500	1	reverse voltage	$1.72 \times 10^{-5}(hr^{-1})$
	Pc=150mW	54	54, 000	0		4.30×10^{-5}
	Pc=200mW	38	37, 500	1	light output	1.04×10^{-4}
moisture resistance	Ta=40° C, RH 95%	61	61, 000	0		3.77×10^{-5}
	Ta=80° C, RH 90%	40	40, 000	0		5.75×10^{-5}
storage	Ta=100° C	56	56, 000	0		4.11×10^{-5}
	Ta=125° C	20	20, 000	0		1.15×10^{-4}
	Ta=-30° C	10	10, 000	0		2.30×10^{-4}
reverse bias	Ta=75° C, V_R+-3. 0V	44	44, 000	0		5.22×10^{-5}
	Ta= 40° C, RH≥95%	44	44, 000	0		5.22×10^{-5}

(2) The results of environmental tests

items	test conditions	sample quantity	failure
temperature cycle	-30° C~90° C, 5cycle	511	0
	-30° C~100° C, 50cycle	83	0
	-55° C~125° C, 50cycle	71	0
boiling	100° C, 20hr	31	0
thermal shock	0° C(1min) 100° C(1min)	94	0
pressure cooker test	120° C, 2atom, 20hr	44	0
moisture resistance	MIL-STD-102B	70	0

INDEX SERIAL NUMBER — 1069

FIELD DATA COLLECTION AND RELIABILITY ASSURANCE OF SOLID STATE DEVICES FOR INDUSTRIAL USE

Yoshihiko Nakamura
Katsuhisa Kimura
Yoshihito Shirai
Omika Works of Hitachi, Ltd.
Hitachi, Japan

Summary

This paper presents actual field data and reliability assurance activity of solid state devices used mainly in a rolling mill plant. In addition, a reliability program based on field data and actual failure rates of transistor type solid state devices and component parts are covered.

Introduction

In recent years, in Japan, the application of process computer control systems in the steel industry is increasing in number; almost all rolling mill plants of today are equipped with process computer control systems.

Hitachi has been applying solid state devices in the control equipment which amplifies and conditions the computer signals to drive process actuators.

Now in the 1970's integrated circuits are being used as a main component of solid state devices, but in the 1960's transistors and magnetic amplifiers comprised the major part of them. In the beginning of 1960's, the use of transistor for the industrial control had been limited to the logic control unit due to the temperature sensitivity of germanium transistor. By the use of silicon transistor many kinds of solid state devices have been developed since 1964. During the last twelve years, approximately 150,000 devices have been manufactured and delivered to customers. The major kinds of these devices are as follows:

1. Operational amplifier circuit unit for analog operation in closed loop control circuit of motor (Fig. 1)

2. Logic circuit unit for logical operation

3. Floating amplifier for insulation of analog control signal

4. Gate pulse generator for shifting the firing angle of thyristor (Fig. 2)

5. Voltage stabilizer as constant voltage sources for operational amplifier circuit unit, logic circuit unit and so on.

These solid state devices have been developed for indistrial use and have mainly been applied for on-line control equipments in rolling mill plants. In many cases, these equipments are installed in a room partitioned on the same floor from a rolling mill yard without air conditioners. Moreover failures of these equipments cause interruption of production. Therefore, they require high reliability under adverse environmental conditions such as high degree of temperature, humidity, dust, gas, vibration, shock, noise and so on. Therefore, reliability assurance based on the actual failure rate data under adverse conditions becomes a prerequisite.

Existing solid state devices delivered from Hitachi have been increasing in number since first delivery of transistor type operational amplifier circuit unit in 1965. As a result of the increasing numbers, we have experienced many troubles in the field. Therefore, we established a reliability program in 1968 on the basis of collecting and analyzing field data in the past three years. All complaint information including infant mortality is being fed back and corrective action is reviewed to monitor product quality.

Reliability Program

Our activities for improving the reliability of solid state devices are shown schematically in Fig. 3.

1. Required failure rate is defined in consideration of total failure rate for the control system and actual failure rate in the field. In 1968, required failure rates for transistor type solid state devices were given. Afterward, the required failure rate for the new device was given at the beginning of development.

2. Estimated failure rate is calculated at the development stage on the basis of actual failure rates of parts. When calculated value is poorer than required failure rate, design for the device is modified to improve reliability.

3. Field data in all the working plants should be collected. Our data gathering methods are as follows:

a. Initial operation test reports: Initial operation test is performed by Hitachi's field engineers. They will report all the failures which occurred during this test period. However it is necessary to reject destroyed devices by misoperation.

b. Replacement and repair: Almost all the failure informations in normal operation are given us by users. Failed devices are replaced by spare devices and returned us so that we may investigate the cause of failure. After analysis of the cause, they are usually repaired and sent back to the user. Sometimes to the user's site we send an engineer who investigates the cause of failure.

c. Maintenance Logbook: As every user keeps a maintenance log, service engineers should call on the user and get the failure informations from the logbook.

4. Actual failure rates of solid state devices are calculated on the basis of the field data. They are used for the reliability prediction of control systems.

5. Actual failure rates of component parts are also derived from the field data. They are used for the reliability prediction of solid state devices, namely the calculation of estimated failure rate.

6. Analysis of the field data is performed by the aid

of the Weibul distribution diagram. By this analysis, defects which require reliability corrective action are revealed, and this information is directed to the appropriate area of responsibility such as design, manufacturing, installation, maintenance, and repair. As a result of this activity, reliability improvement will be obtained and the actual failure rate will be made less than required failure rate for each device.

Since 1968, we have been performing reliability improvement activity according to the reliability program described in the preceeding statements, so that the reliability of solid state devices has been improved remarkably. In the following statements, we would like to show the reliability improvement activity based on the actual field data.

Data Collection in Initial Operation

In general, installation of electrical equipments in computer control systems is divided into two groups according to time sequence. The first group consists of usual control equipments, electrical equipments, electro-mechanical equipments and so on. The second one consists of control computers and their peripheral equipments.

After the first group equipments have been installed and tuned individually, initial operation of of the control system is performed without a control computer. An initial operation test is performed by Hitachi's field engineers. When the system includes newly developed equipments or solid state devices, the initial operation test has an especially important meaning. For example, sometimes we experienced defects of parts, problems of noise and distortion of waveforms against gate pulse generators and floating amplifiers and so forth.

Fig. 4 shows accumulated failure distribution diagram of actual rolling mill plant "A". Logarithmic scale is used for both ordinate and abscissa. Solid line shows accumulated number of failures for the total system which contains all of the electrical equipments such as motor, selsyn, magnetic valve contactor, limit switch, thyristor, solid state device and so forth. Dotted line shows accumulated number of failures for transistor type operational amplifier circuit unit included in the system.

The meaning that this diagram designates is as follows:

1. The line with an angle of 45 degree means constant MTBF (Mean Time Between Failures).

2. The line with a larger angle than 45 degrees against horizontal line means that the failure rate is increasing. (MTBF is decreasing.)

3. The line with a smaller angle than 45 degrees means that the failure rate is decreasing.

This diagram is effective for us so as to catch the reliability status of a total system.

Plant "A" in Fig. 4 has an extremely large scale control system including more than 7,000 solid state devices.

The solid line in Fig. 4 indicated to us that the failure rate was increasing after 30 days. (Because the angle against horizontal line is larger than 45 degree.) Therefore, the Pareto's Diagram was described as Fig. 5. As a result of analysis by Fig. 5, it was shown that the cause was due to failures of transistor type operational amplifier circuit unit. So in Fig. 4, the dotted line showing accumulated number of failures for transistor type operational amplifier circuit unit was added.

At the same time, several numbers of failed units were returned to our factory and their cause was investigated. At last, it was revealed that the cause was a lot failure of a specific kind of transistor. In the production process, several kinds of screening tests were performed, but these defective transistors would not be rejected. For the plant "A", we substituted all of that type of transistor by newly manufactured ones. For purchasing specifications of the transistor, a few items were added to prevent the same kind of troubles.

Fig. 6 indicates an example of plant "B" of which the control system includes about 3,500 solid state devices. Solid line shows accumulated number of failures for the total system including motor, contactor, thyristor, solid state device and so on. As a result of reliability improvement activity previously mentioned, failures of solid state devices decreased extremely. The reliability status of plant "B" became much better than plant "A".

In initial operation tests several years ago, we experienced unpredictable lot failures of component parts and defects in design due to lack of experience. But now in the 1970's, initial failures of solid state devices such as used in plant "A" are rarely observed because of reliability improvement of component parts and improved company engineering capability.

Field Data Analysis

Now, we would like to analyze the field data for each kind of solid state devices individually. In the preceeding chapter, we explained about plant "A" and "B" which have very large scale control systems. But in general, most of the plants have smaller control systems which contain several tens of operational amplifier circuit units. It is not only troublesome that field data for each plant are analyzed individually, but also unreasonable in statistical analysis because of small population.

Therefore, we decided to consider devices manufactured in every six months as one lot for each kind. For example, lots in 1972 are divided into two groups as follows:

1972 (1) lot: Mar. 21, 1972 - Sept. 20, 1972
1972 (2) lot: Sept.21, 1972 - Mar. 20, 1973

The reliability status of each kind of solid state devices is described on the weibull distribution diagram. Failures of several lots are plotted on the same diagram for a long period so that an abnormal pattern of a specific lot is easily recognized.

Fig. 7 indicates Weibull distribution diagram of transistor type operational amplifier. Long dotted line, indicating 1967 (1) lot, gives us standard pattern. On the other hand solid line, indicating 1966 (2) lot, goes up abruptly around 30 months. The cause of this failures was analyzed and disclosed to be broken film in carbon film resistors. The failure rate for carbon film resistor in 1966 (2) lot was plotted on the same diagram as short dotted line. This diagram gave us the information as follows:

1. Shape parameter "m" for transistor type operational amplifier circuit unit:

1966 (2) lot: m = 4.1
1967 (1) lot: m = 0.3

2. Shape parameter "m" for carbon film resistor:

1966 (2) lot: m = 6.2

We decided failures in 1966 (2) lot as wearing out pattern according to this information. Therefore, we substituted all of the transistor type operational circuit units in 1966 (2) lot by new ones.

In this manner, we are performing reliability assurance activity by means of surveying failures of solid state devices in actual fields.

Actual Failure Rates

Failure rates calculated on the basis of field data which have been collected since 1965 are shown in Table 1 for solid state devices and in Table 2 for component parts.

Major solid state devices using silicon transistor are listed in Table 1. In the table, required failure rates were decided in consideration of total failure rate for control system in 1968. First of all, failure rate for logic element (plug-in unit) was given as 10^{-7} per hour. Then failure rates for the other devices were decided by comparing with logic element. Failure rates in column (A) designate the result of calculation for all of the solid state devices manufactured until March, 1972. On the other hand, failure rates in column (B) show the result of calculation for the solid state devices manufactured after 1969. We had established the reliability program above mentioned in 1968 so that the failure rates for lots manufactured after 1969 are distinctly under the required value.

Failure rates in column (A) are used for the reliability prediction of control systems.

Failure rates for major component parts are listed in Table 2. They are used for the reliability prediction of solid state devices.

Conclusion

The reliability program based on actual field data was established in 1968. We have been performing reliability improvement activity according to this reliability program since 1968, so that actual failure rates for solid state devices manufactured after 1969 distinctly satisfy the required failure rates.

In 1970, we started to apply a new series of solid state devices of which the main component parts are Integrated Circuits, for industrial use. We have already manufactured several thousands of operational amplifier circuit units, but only two failures have been experienced in the actual fields during the last two years. So we are confident the new series of sodid state devices have extremely high reliability.

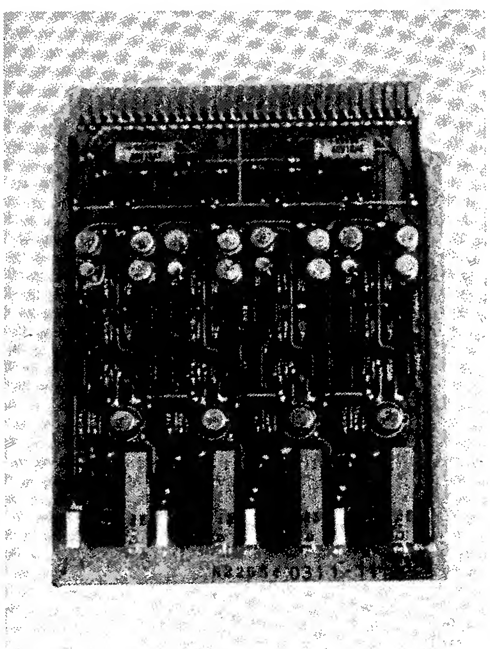


Figure 1 Operational Amplifier Circuit Unit

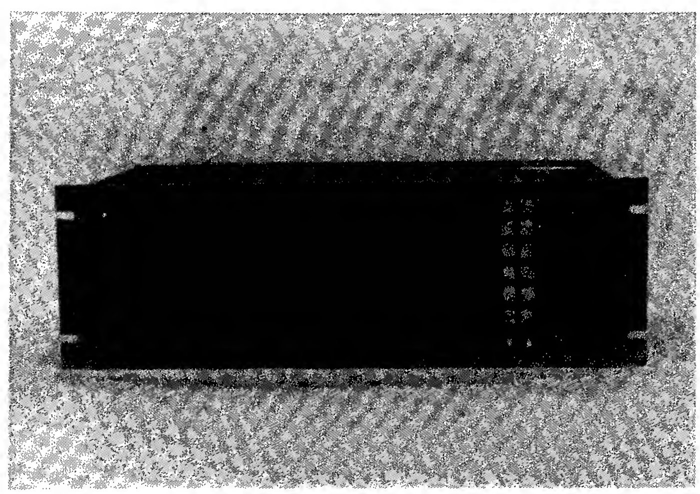


Figure 2 Gate Pulse Generator for Thyristor

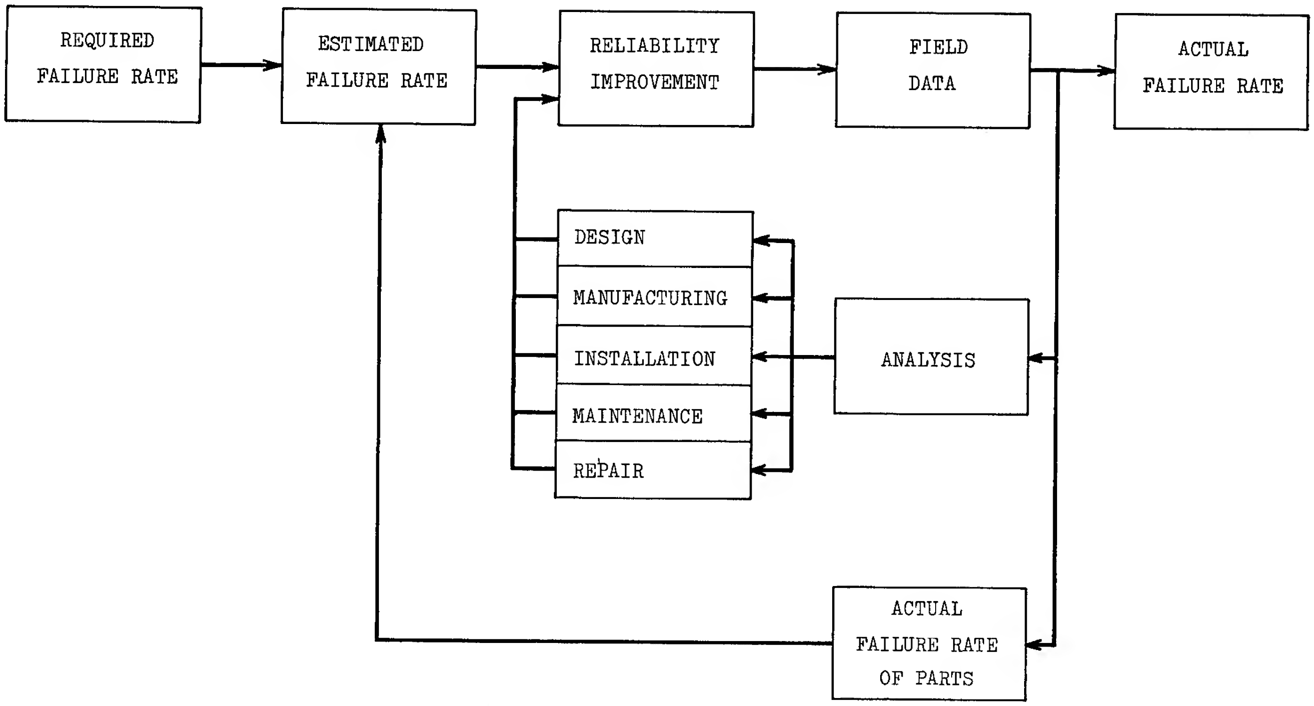


Figure 3 Flow Chart of Reliability Program based on Field Data

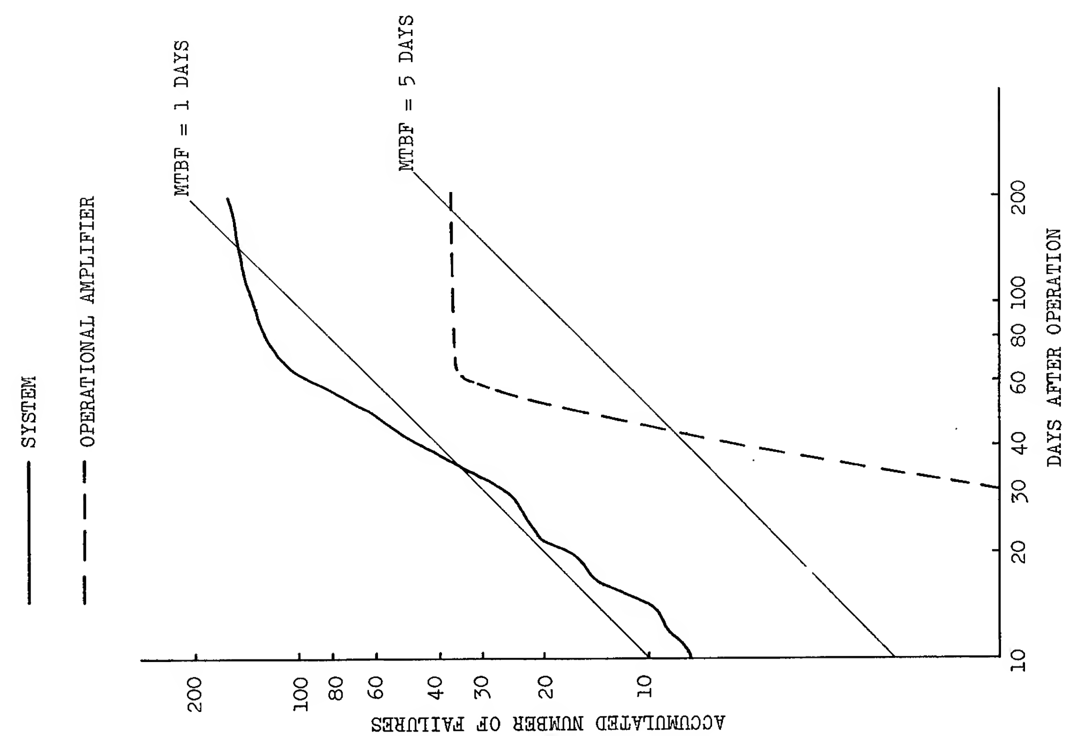


Figure 4 Accumulated Failure Distribution Diagram for Plant "A"

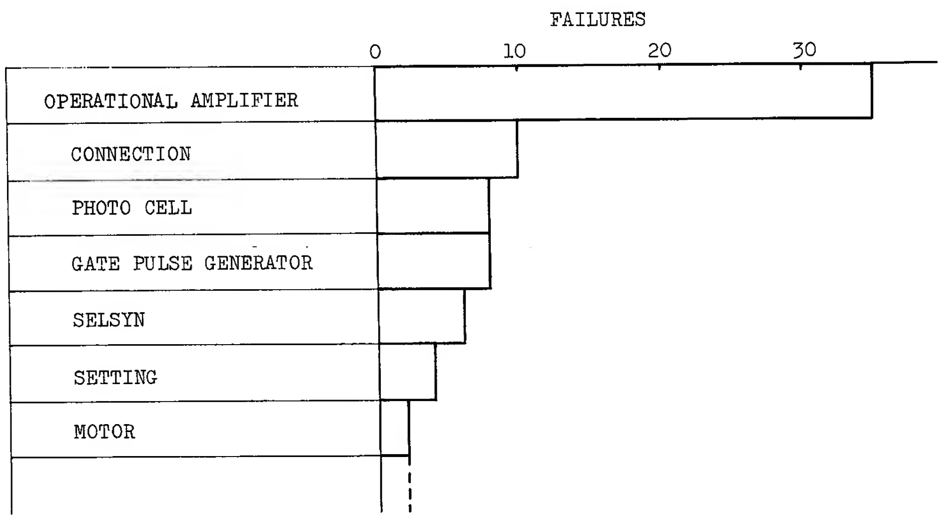


Figure 5 Pareto's Diagram for Analysis of Failures

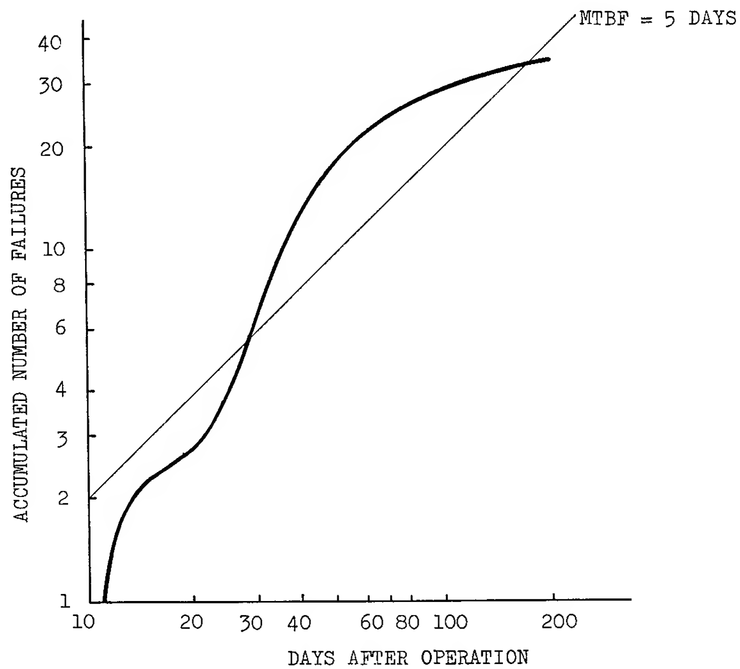


Figure 6 Accumulated Failure Distribution Diagram for Plant "B"

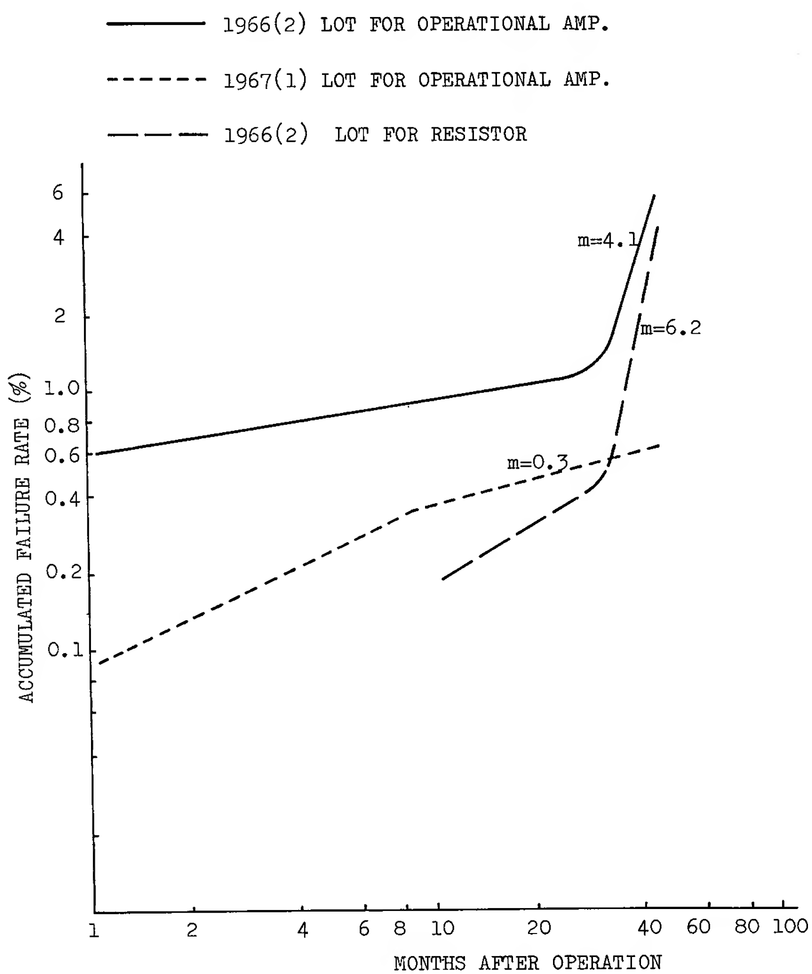


Figure 7 Weibull Distribution Diagram for Operational Amplifier Circuit Unit

Table 1 Actual Failure Rates for Solid State Devices

DEVICE \ FAILURE RATE	REQUIRED F.R. (10^{-6}/Hr)	(A) TOTAL F.R.(10^{-6}/Hr)	(B) IMPROVED F.R.(10^{-6}/Hr)
OPERATIONAL AMPLIFIER	0.2	0.37	0.16
GATE PULSE GENERATOR (A)	2.0	2.2	1.8
GATE PULSE GENERATOR (B)	0.2	0.42	0.19
FLOATING AMPLIFIER	0.1	0.08	NO FAILURE
LOGIC ELEMENT	0.1	0.07	0.09
VOLTAGE STABILIZER (A)	0.3	0.26	NO FAILURE
VOLTAGE STABILIZER (B)	0.3	1.4	NO FAILURE
VOLTAGE STABILIZER (C)	0.8	1.3	0.71

Table 2 Actual Failure Rates for Component Parts

PARTS		FAILURE RATES (FITS)
SEMICONDUCTOR	SWITCHING TRANSISTOR	4.1
	TWIN TRANSISTOR	41
	POWER TRANSISTOR	16
	UNIJUNCTION TRANSISTOR	83
	DIODE (for LOGIC)	0.08
	DIODE (for RECTIFIER)	1.7
	ZENER DIODE	11
RESISTOR	FIXED CARBON FILM	1.4
	FIXED WIRE	4.0
	VARIABLE CARBON FILM	2.1
	VARIABLE WIRE	27
CONDENSER	ELECTROLYTIC	NO FAILURE
	METALLIZED PAPER	1.5

COMPONENT DEFECTS AND SYSTEM RELIABILITY

G. D. Goldshine and R. E. Martin
General Dynamics, Pomona Division
Pomona, California

SUMMARY

Apparently random failures in final test of a tactical weapon system were traced to incipient defects in electronic piece parts. An interim program to minimize these failures was developed which included additional screening of semiconductor devices and other part types and subassembly conditioning (burn-in and vibration). The major cause of incipient defective parts was found to be part supplier workmanship. A comprehensive program has been developed to eliminate this failure made as a significant factor in future programs.

Introduction

The primary product line of the Pomona Division of General Dynamics consists of tactical weapon systems. This paper is concerned with one particular system which has been in production for several years. It is a moderately complex system having, in addition to the usual mechanical and electromechanical parts, electronic circuitry totaling about 6000 parts. Construction of the electronic section is fairly conventional, consisting of cordwood modules (Chart 1) and boards (Chart 2). Interconnects are primarily soldered, although a few welded assemblies are utilized from earlier versions of the system.

About two years ago, a modified version of this system was put into production. It represented a significant advance in performance and a corresponding increase in complexity. The quality requirement, by contract, was high. A success rate demonstration (SRD) of 96% was specified. The success rate demonstration is performed at the section level and uses automatic checkout equipment to test a large number of parameters which verify individually and/or in operating combination the acceptability of the elements in the section. Thus, each parameter tested represents an opportunity for failure. The predominant mode observed is only one parameter failure in any failed section. In this predominant mode, these acceptance criteria demand a very high inputted reliability (Chart 3). This meant that each month's production of each section, when presented for final acceptance testing, could have no more than 4% rejects. There have been significant problems in meeting this requirement. The basic reason for section failures, once the program had matured, was found to be electronics parts, primarily semiconductors, which were known to be operating satisfactorily when tested at the lower assembly level and initially at the acceptance level but then failed when the section was being tested a second time for acceptance.

The distribution of parts population is shown in Chart 4. As is indicated, the technologically more advanced classes of parts, transistors, and integrated circuits comprise 10% and 3% of the population respectively. However, as will be shown, the problems encountered are heavily influenced by this portion of the population.

Problem Description

Now that we have seen something of the background, it would appear that the problem is basically that electronic parts are failing, catastrophically in most cases, in situations where there should be few or no failures, in top level assemblies (sections) where repair is most expensive. Chart 5 shows the distribution of part failures for a typical month. As can be seen, the failures are heavily weighted toward the semiconductor devices, with capacitors forming the major portion of the remainder. This distribution is for total failure at all levels. The failure distribution by part class is approximately the same at each level of assembly and test. Chart 6 shows the distribution of failures by assembly level. As might be expected, the number of failures is highest at the lowest level of assembly and decreases as the modules are assembled into plates and plates into sections. Each level of assembly, incidentally, represents a number of tests with fail-fix operations after each test as required. So, why should a component which has gone through an extensive series of tests, as a component and as a part of an assembly, suddenly fail? The investigation into this phenomenon and actions to minimize it in the current program and in future programs form the basis of this paper.

Chart 7 shows distribution of failures by mode and time. As can be seen, during the initial phases of the program, a good mix of failure modes was observed. Workmanship external to the component was significant, as were overstressed parts (ZAP's) and parts which were thought to be bad when removed, but which, on diagnosis, were found to be good. As the program matured, these modes decreased radically and the major cause of failure at top assembly tests was found to be electronic parts, and the mode of failure was internal supplier workmanship. While this poor workmanship took several forms, the two predominant modes were (1) bad internal connections and (2) contamination by conductive particles. It is interesting to note at this point the failure modes which did not occur at top assembly tests. Out-of-limit parts, shorted capacitors, open resistors, semiconductors in which the junction was defective, and similar failure modes which might be expected of marginally weak parts operated under significant electrical stress just did not occur. This is not to say that marginally weak parts are not manufactured as part of the lots which are delivered to Pomona Division, but apparently the stress screens and tests applied at the part level and the several assembly levels are effective in removing them prior to top assembly testing. Chart 8 shows the distribution by part class of top assembly (section) part failures. Chart 9 shows frequency of incidence by part type. As can be seen, while a few parts can be considered modal, the majority of failures occurs in small numbers (1 and 2 failures in the entire program) and must be considered truly random.

Let us consider the components we have been discussing. They are largely conventional electronic parts, basically off-the-shelf items, which are, however, carefully defined and have extensive characteristic and quality control invoked. At the time of the basic system design (1964-65) it was deemed that existing military specifications were not sufficiently rigorous, and specification control drawings for all part types used were created. These call out, for all transistors and integrated circuits, and for most other types,

100-percent process conditioning, including power burn-in (or reverse bias at high temperature) and other thermal and mechanical stresses to weed out the weaker "infant mortality" parts. Sample tests for all electrical and environmental parameters and a life test are included. These screens and tests are performed by the supplier. Receiving inspection and environmental tests are performed in-house on a sample basis to verify the conformance of each received lot to the specification requirement. Typical requirements are given in Chart 10. During the design phase all circuits were subjected to a rigorous analysis for electrical stress to ensure that the derating factors in General Dynamics design criteria were observed. Environmental factors of the missile application, temperature, shock, vibration, and so forth, were similarly analyzed to ensure that a suitable safety margin was included in the part specification.

So, we find ourselves with good parts, screened and tested, sample tested, sample tested again, put into application which have been checked for electrical and environmental stress, tested again in the application, at several levels of assembly, and then a very small but unacceptable percentage of the parts continues to fail at the critical final test of the top assembly.

When we talk of very small but damaging numbers, it is just that. Half a dozen failures a month, out of a parts population of close to half a million parts per month, or about .001 percent failure rate, is by normal standards very good performance. However, looking at the part categories in which failures are observed, primarily transistor and integrated circuits, with a combined missile population of about 700, the percentage is still small, but now becomes more significant. The important factor, though, is that failure of a single part may affect the nominal missile performance. Since the final top assembly test is considered to be an approximate dress rehearsal of the performance of the system, this does not mean the operational performance will be degraded. Any number of part failures, however small, is certainly a matter of concern and worth a good deal of effort to correct.

When the failed parts were examined in detail, the two failure modes previously mentioned were found to be predominant, defective bonds and conductive particle contamination. Microphotographs of typical cases of both are shown in Charts 11 and 12. These two failure modes were observed in a large number of part types, but the mechanism was certainly modal - a deficiency in the processing of canned semiconductor devices.

Attempts at Solution

Having to some extent zeroed in on the problem, the search for appropriate corrective action began. The first step, on which action could be taken immediately, was to determine which part lots were suspect, based on the premise that delivered lots were fairly homogeneous and that if several failures of a specific mode were observed, there were probably more potential failures. Lacking any effective screen at the time, these lots were scrapped. This certainly cost us many good parts, but it was felt that this was preferable to allowing any parts with incipient defects to be built into the missile. Next, or actually concurrently, the suppliers were called in, to make them aware of the problem and to have them initiate action to correct the deficiencies in their production processes. This was only partially successful. All suppliers were cooperative and listened attentively to our problems. Specific corrective action, however, was something else. The failure percentage we were seeing was, on an absolute basis, so low that even though the suppliers agreed that the specific parts which were shown them, as in the slides you have just seen were defective, they could not feel that a major effort to improve their basic product was warranted. In conjunction with the suppliers the amorphous world of "Hi-Rel" parts was explored and the benefits of pre-cap visual examination, captive lines, and other classic Hi-Rel operations were discussed. These certainly gave some promise of decreasing the incidence of failures due to poor workmanship, but raised enough questions of part availability and economics that nothing specific could be done at the time.

This situation left us with the necessity of developing screens which could be used to weed out incipient defectives in existing lots or newly manufactured lots which presumably would be not much different from what we had been receiving. Consulting once again with the suppliers, some things were discussed and tried. Defective bonds could, perhaps, be made to open completely with high acceleration shock. A 30-KG shock test was tried and found wanting. This screen did weed out some parts, presumably those with weak bonds, about 5 percent of the parts in the lots screened. This looked encouraging. However, when the survivors of the lot were assembled into system applications, more failures, still for bad bonds, were observed at the top assembly level tests. Whether the percentage of failures had decreased was hard to tell, considering the small number involved, but shock testing was certainly not the complete answer. Centrifuge tests gave similar results. Gold bond wires had a small, but significant force applied to them by 30-to-40-KG acceleration, enough perhaps to separate an already nonexistent bond, but not enough to part a weak one. The force applied to aluminum bond wires was found to be negligible. In some part types, the supplier has agreed to change his actual device design, to use heavier lead wire, and thus, presumably, obtain stronger bond, but whether it significantly reduces the already small number of incipient defective bonds remains to be seen. On other part types, the occasional bond failure is still with us. Sample bond pull test on a relatively large sample will tell whether the lot has some number of bad bonds in it, but this is an "after the fact" test and suitable only for acceptance of lots.

Temperature cycling, which would hopefully exercise weak bonds to the point where they would fail, was attempted. The results again were inconclusive. In some lots a significant number of devices developed open bonds, but so far there is insufficient data to show that the survivors are all good.

Screening for particle contamination seemed at first to be an almost impossible task, and early attempts to use either X-ray or a short circuit indicator as a particle detector while vibrating the part were ineffective. However, the acoustic detector equipment, which detects the noise of a "rattling" particle while the part is vibrated, has proved to be extremely effective. Too effective in one way - it will pick up noise generated by any internal particle, whether it is conductive or not. This results in some number of parts with glass particles being rejected, but this is a low price to pay for the assurance that parts with potentially disastrous metal particles are being screened out of the system.

While the major problem observed has involved transistors and integrated circuits, capacitors have also given a significant amount of trouble, as was shown in one of the earlier charts. The predominant failure mode seen was bad internal connections, from the capacitor element to the lead wire. Fortunately, this was easier to handle than the semiconductor problem. The suppliers seemed startled when we showed them X-rays and cross sections of their devices,

but recognized the problem and were able to correct their assembly processes fairly quickly. X-ray and tap tests provided adequate screens for parts already in stock.

While we were struggling with the problem at the part level, another approach, that of further exercising the assemblies, was explored. We have seen that repeated testing of assemblies does screen out many defective parts. Would additional testing, or burn-in, at a selected assembly level provide the required additional screening? An experiment was set up to exercise plates by subjecting them to a preconditioning cycle consisting of functional burn-in for 10 hours, followed by 45 to 60 minutes of vibration with power applied. It was hoped that this would accelerate the failure of any marginally weak parts. The results of this exercise were as expected. The failure rate closely approximated that which non-preconditioned plates exhibited when assembled into sections and tested at that level. This screening process has helped significantly in causing incipient failures to occur at the plate level, where repair costs are lower than at the initial test at section level. (See Chart 13.) However, the fact that this screening process is only partially effective in reducing the small number of failures at the final section test indicates that this is not a final solution to the problem. This approach has been incorporated in the standard processing of the plates.

In the interest of sharing common data and experience in the general area of parts problems, a survey was made of aerospace system contractors. This survey indicated an awareness of and an intense interest in the part problem on the part of all companies surveyed and culminated in a seminar which was held at Pomona in November 1971. Participating companies included:

Aeroneutronics, Newport Beach
General Electric, Utica
Hughes Aircraft, Tucson
Lockheed, Sunnyvale
Martin, Marietta, Denver
North American Rockwell, Columbus
Texas Instrument Company, Dallas
Stromberg-Carlson, Rochester
General Dynamics, Convair, Forth Worth Operation
General Dynamics, Convair, San Diego Operation
General Dynamics, Electronics Division, San Diego
General Dynamics, Pomona Division, Pomona
General Dynamics, Electronic Division, Orlando

A summary of the results of this seminar is given in Charts 14 and 15. As can be seen, part quality is considered to be a general problem throughout the industry. Various testing and screening programs have been devised to elevate part quality. These are costly and not entirely effective. Apparently, further, more effective controls are needed to achieve the quality needed for newer and more advanced weapons systems.

Future Programs

Thus far we have discussed the discovery of a problem and an investigation into corrective action. Some of the corrective actions could be, and were, put into effect immediately. However, considerable time was needed to develop and implement a fully integrated program to take advantage of all we have learned. This program, which is shown in Chart 16, selectively applies screens and tests to part types which have had a history of failures or which, based upon construction and supplier history, seemed to be good candidates for future trouble. This program will be in effect during the remaining production years of the present design.

This must still be considered an interim program. It is based on the present design, which, of course, includes the present part selection and present part design in terms of the specification control drawings. The vital consideration is to apply the experience, good and bad, which we have had to the design and the quality parameters of the next generation.

This is being done in the design of a follow-on system. As might be expected, this system is functionally far more complex than its predecessor, with a much higher part density and greater use of complex (MSI and LSI) devices and circuits. Requirements for greater precision in operating characteristics and operability over a greater range of environments reemphasize the need for added attention to the quality/reliability problem. Starting from the beginning, every effort is being made to ensure that part selection, fabrication, screening, and testing are such that parts as assembled into the system are in fact as well as in name "high reliability." This will be, however, in the context of the system application, with its requirement of a very high probability of satisfactory performance for a relatively short mission, rather than the classical hi-rel concept of long life.

Whenever possible, military high-reliability parts (ER, TX, TXV, and MIL-M-38510 level A or B) are used. In some cases, even these specifications do not completely meet our needs, and additional tests or screens, some to be performed by the supplier and some in-house, must be imposed. A certain amount of caution must be observed, however. These specifications are intended to represent the best parts generally available. However, in many cases, particularly microcircuits specified in MIL-M-38510, fully qualified suppliers for a reasonably wide range of part types are not now available. It is to be hoped that time will rectify this situation. Part procurement documentation prepared at General Dynamics for parts which are not available as military hi-rel, will follow the intent of the military hi-rel specifications, with, of course, any additional tests and screens which are required.

In addition to the procurement specification, which defines the characteristics and quality of the part and the tests required to verify then, it is now recognized that a positive effort must be made to ensure that the quality required is inherent in the design and fabrication process of the part. While our relations with our part suppliers have always been close, any detailed study of their process has occurred only in the case of serious trouble, because of a natural reluctance to reveal processes considered proprietary. In the finalization of the part complement for the follow-on system, studies will be made of the part design and fabrication process to ensure that the part has adequate inherent resistance to the failure modes which have caused problems in the past. In-process inspection, such as pre-cap visual inspection and large sample bond pull tests, will be invoked as required. It is felt that a cooperative effort with the part suppliers, with well defined goals and good management at both ends will do a great deal to minimize problems in parts as received. In-house sample testing on a regular basis will provide an audit on the effectiveness of the supplier's efforts.

Application of parts is being carefully controlled. An active standardization program will reduce to a workable minimum the number of part types and permit greater attention and more detailed testing to be given to each lot of each part type. A detailed computer-aided analysis of each critical circuit is being made to ensure that electrical stress (steady state and transient) conforms to the derating factors established as design guide lines for all operating

environments. To the greatest extent feasible, redundant circuitry is being designed in, to minimize the operational impact of part failures.

The limited experience obtained in the present program with assembly stressing to weed out defectives indicates that this will be a useful tool in the follow-on program. It is planned to provide for combined electrical burn-in and environmental (thermal and mechanical) stress at all levels from first assembly (module or board) through next assembly (plate) to top assembly (section). All parts will of course, be process conditioned, including electrical burn-in and exposure to a series of environments. During the prototype stage in this program a number of assemblies will be subjected to a series of overstress tests to determine most probable failure modes. Based upon the results of the prototype tests and current failure rate data, assembly stress screening will be imposed on an as-required basis.

Failure data, maintained on a real-time basis and available quickly in any format desired, is an essential part of this quality program. With all of the possible precautions taken, there will still be problems, which must be discovered and identified quickly if timely corrective action is to be taken. A consolidated computer data bank, with all quality information inputted, will provide the basis of a fast feedback system. All data relative to a particular failure will be available, including supplier process and lot acceptance test data, in-house test data, performance of the part in other assemblies, yield history of the assembly, performance history of the assembly in its next higher assembly, prior failure modes, and so on.

A similar integrated approach will be taken to the test program to ensure that each test, at each level, is a true evaluation of the capability of the item tested to perform properly in its next higher assembly. In this as in any program, the economic factors, in terms of rework and scrappage cost at all levels of assembly must be considered; always, however, within the context of the reliability/quality requirements of the end product. A comprehensive study of production yield factors and cost factors has been initiated to provide a means of obtaining production yield and cost targets which will be used by designers as guidelines in their design. This study will relate the factors which affect production yield, primarily at the lowest level of assembly, but also at the higher levels. These factors include the quality of the parts used, how forgiving the design is in terms of part parameter variation, and the ability of the test program to evaluate each assembly properly in terms of next assembly requirements. A similar analysis of cost factors will determine the optimum allocation of effort (and money) among the several levels of assembly. Part quality, expressed as parameter variability as well as catastrophic failure rate, is obviously an important factor in any such analysis. The results of this study will be fed back into the part specification quality requirements and into the in-house screening requirements.

Conclusion

The experience obtained in the design and production of the present system indicates that the attainment of maximum system reliability is primarily influenced by the quality of the electronic parts used. The quality factors involved are concerned with the residual defectives after all normal screening and testing and are primarily due to supplier workmanship.

An extensive investigation has produced additional screens and tests at the part level and at assembly levels which, in the context of an existing design, reduce the incidence of, but do not completely eliminate, part failures in the completed system.

Applying this experience to a new design, the following steps are being taken to minimize or eliminate the problem.

1. Ensure that application, electronic and mechanical, of parts is well within their capabilities, with adequate safety margin.

2. Provide specification for procurement of hi-rel electronic parts based upon system needs and including, on an audit or screen basis, tests for the specific failure modes which have been generally observed.

3. An organized high stress screen and test program at the assembly level to ensure maximum quality of each assembly as it goes into its application.

4. Establishment of an integrated test and data program to provide real time visibility of the quality status of all parts and assemblies at all times and provide predictions of possible future problems.

5. Study of all phases of the production process to ensure that the optimum tradeoff of part quality and overall cost is made, keeping in mind the reliability requirements of the end product.

Above all, an attitude and an approach to high quality must be engendered in all personnel connected with the program. Administration and communication lines are being established to ensure that positive preventive action is taken and, when trouble does occur, that rapid and effective corrective action is taken. Production of any complex item is a dynamic process, and continuing, competent attention to all aspects of the design and production program relating to system reliability is an absolute necessity.

SECTION	NO. OF PARAM TESTED	SECTIONS PER LOT	NO. OF FAILURE POSSIBILITIES	ALLOWABLE FAILURES*	RELIABILITY
GUIDANCE	193	75	14,475	3	0.9997928
CONTROL	200	75	15,000	3	0.9998111
ORDNANCE	57	75	4,275	3	0.9992983

*100% MINUS 96% SUCCESS RATE REQUIREMENT X 75 = 3

F90121 7102

CHART 3 – SECTION SUCCESS RATE DEMONSTRATION

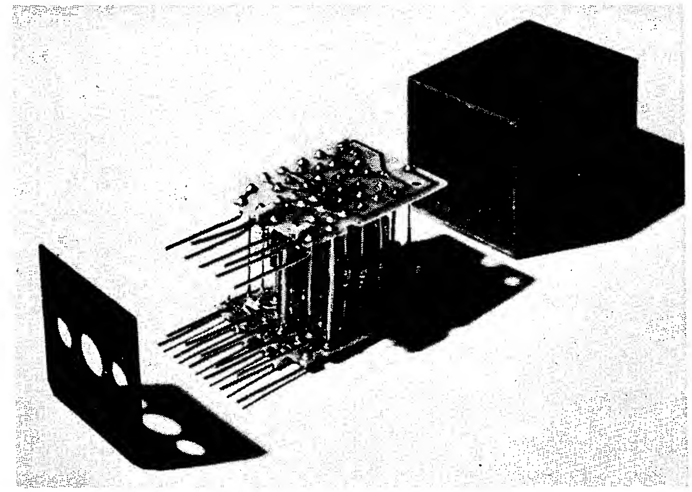


CHART 1 TYPICAL 3d MODULE

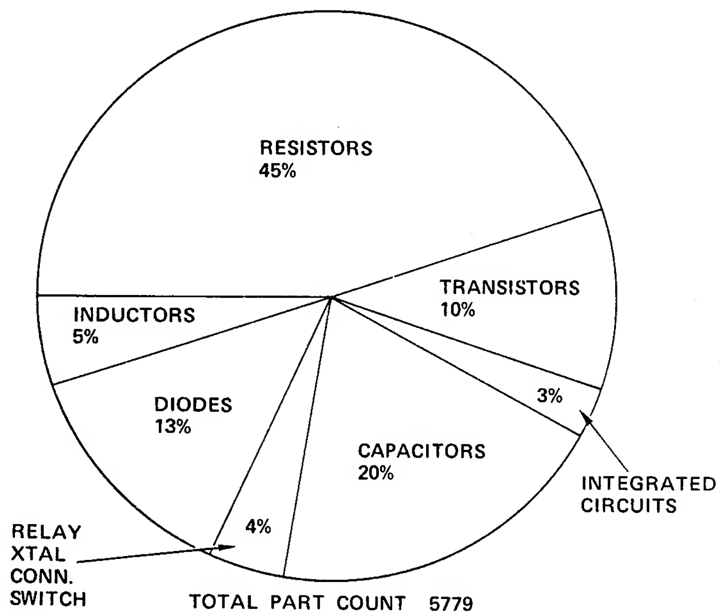


CHART 4 – PART DISTRIBUTION

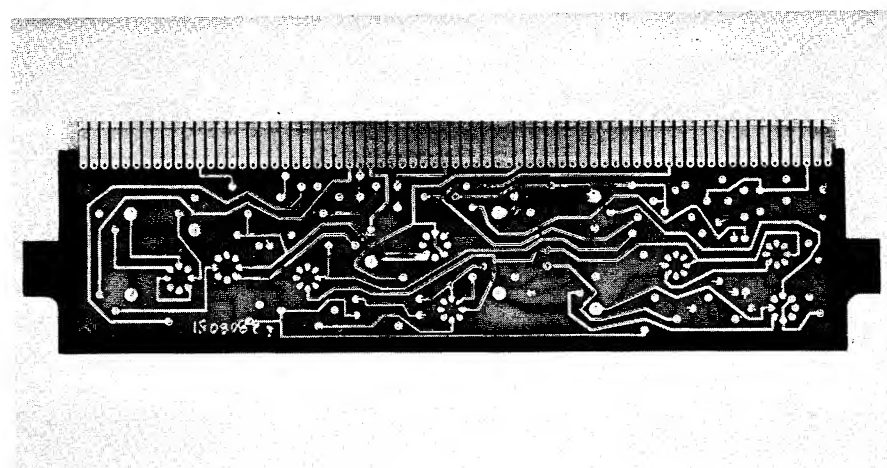


CHART 2 2-d ELECTRONIC SUBASSEMBLY

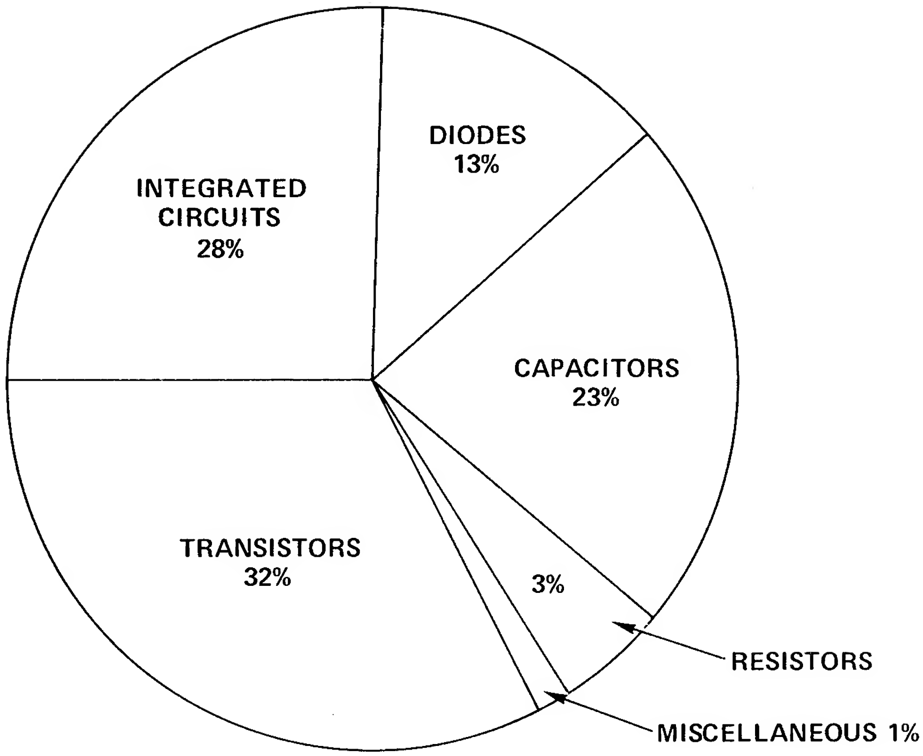


CHART 5 – PROBLEM DISTRIBUTION
TYPICAL MONTH

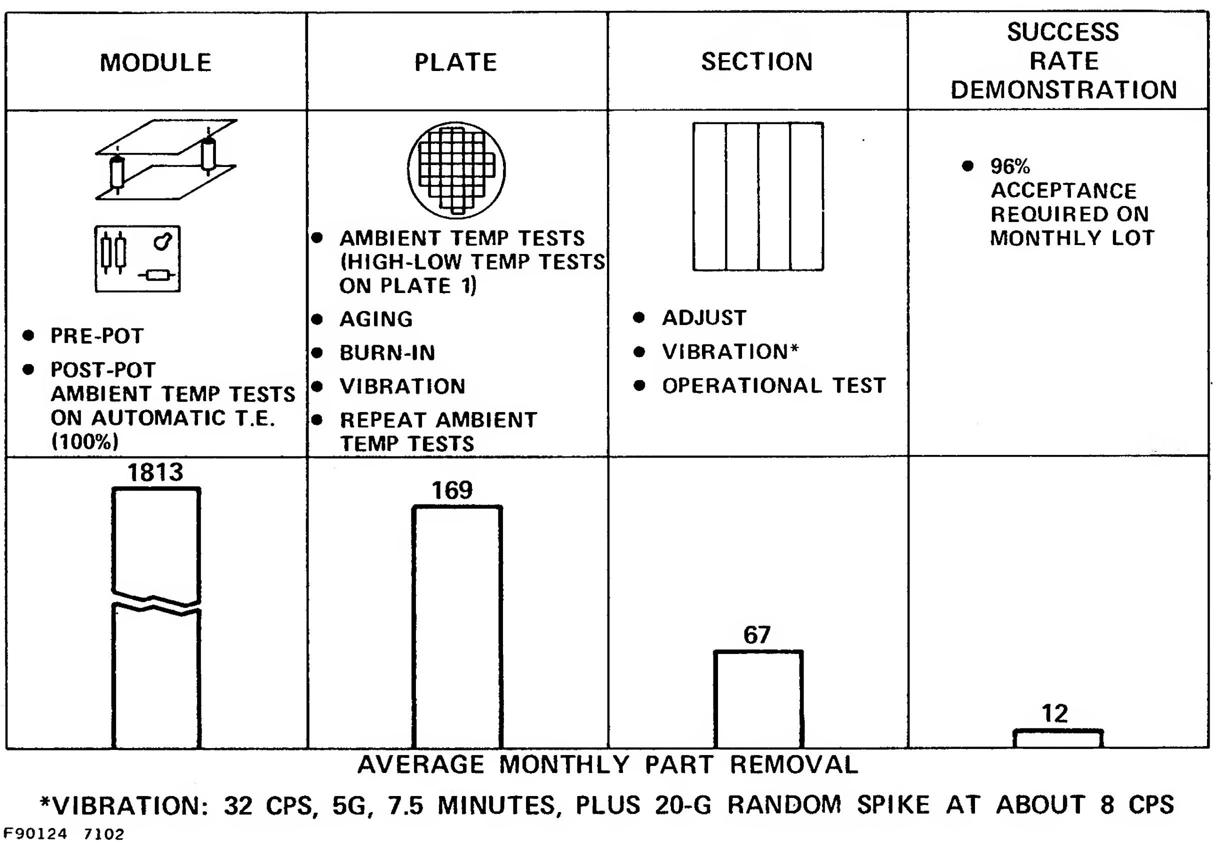


CHART 6 – PART QUALITY SUMMARY – ASSEMBLY LEVEL

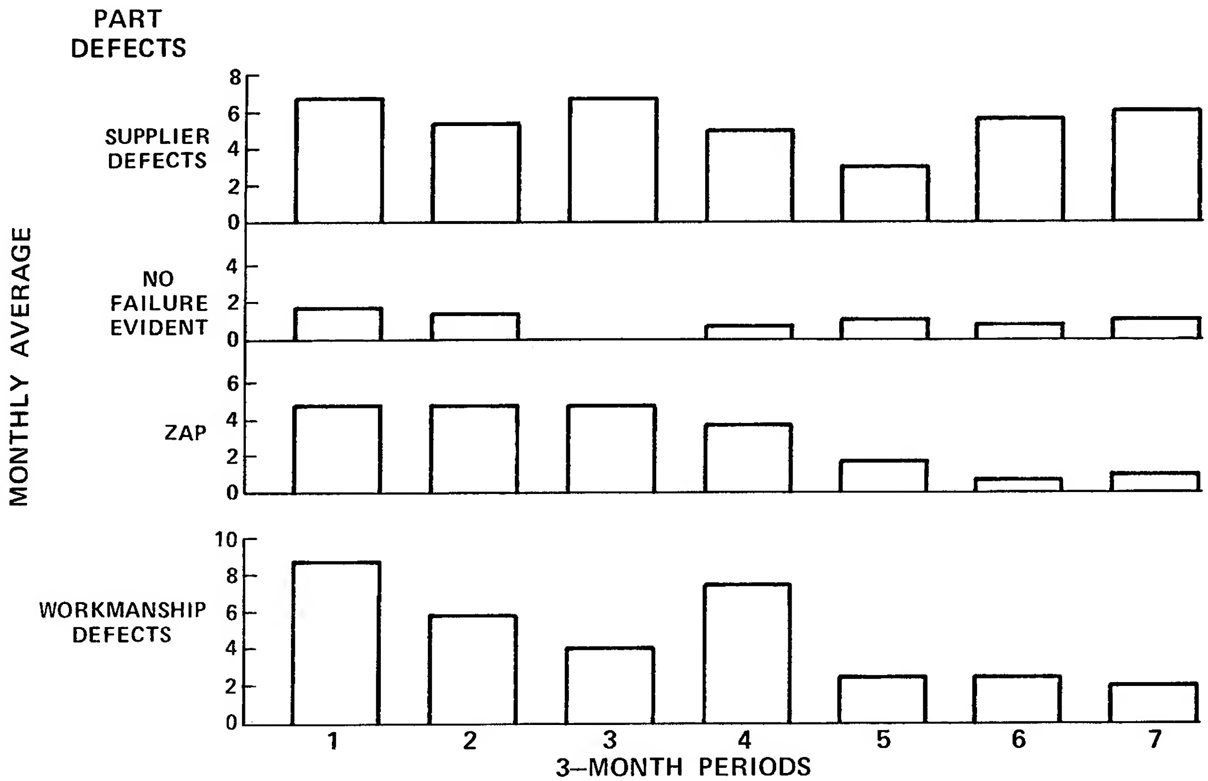


CHART 7 – SRD FAILURES BY DEFECT MODE

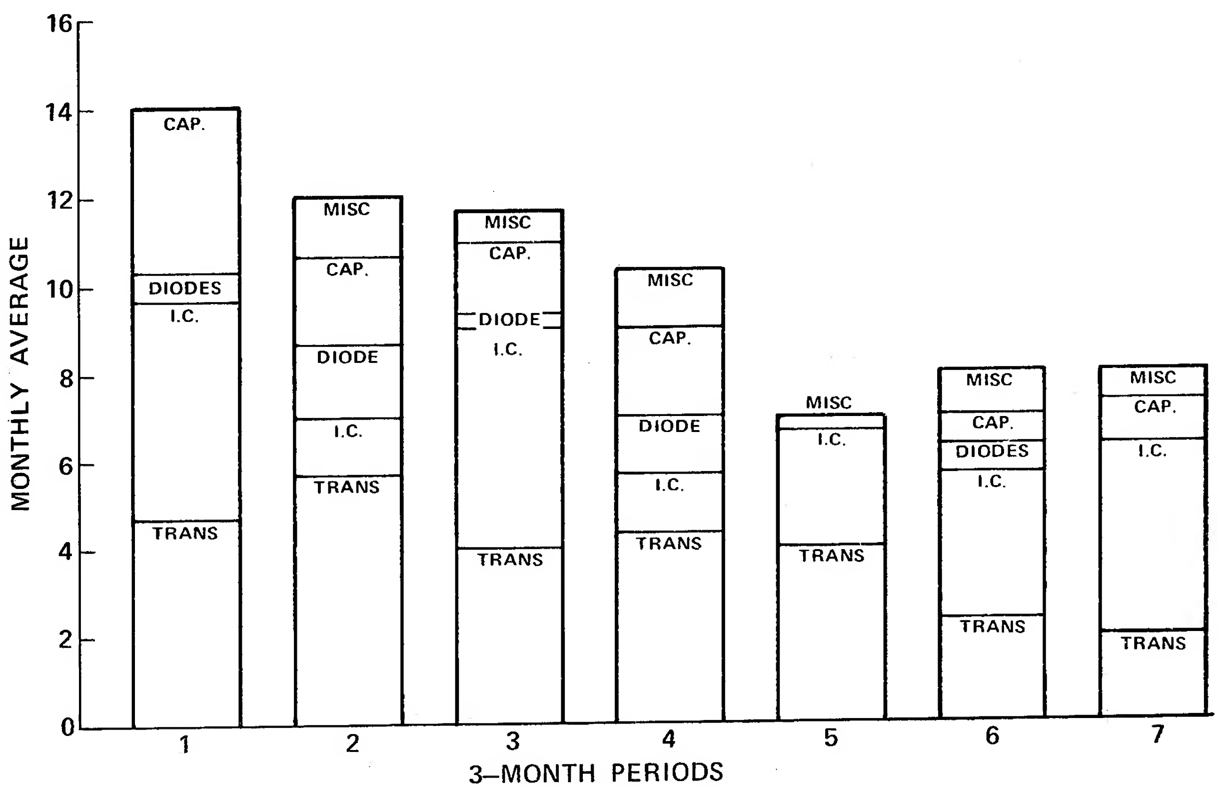


CHART 8 – PART REMOVAL HISTORY BY PART CATEGORY

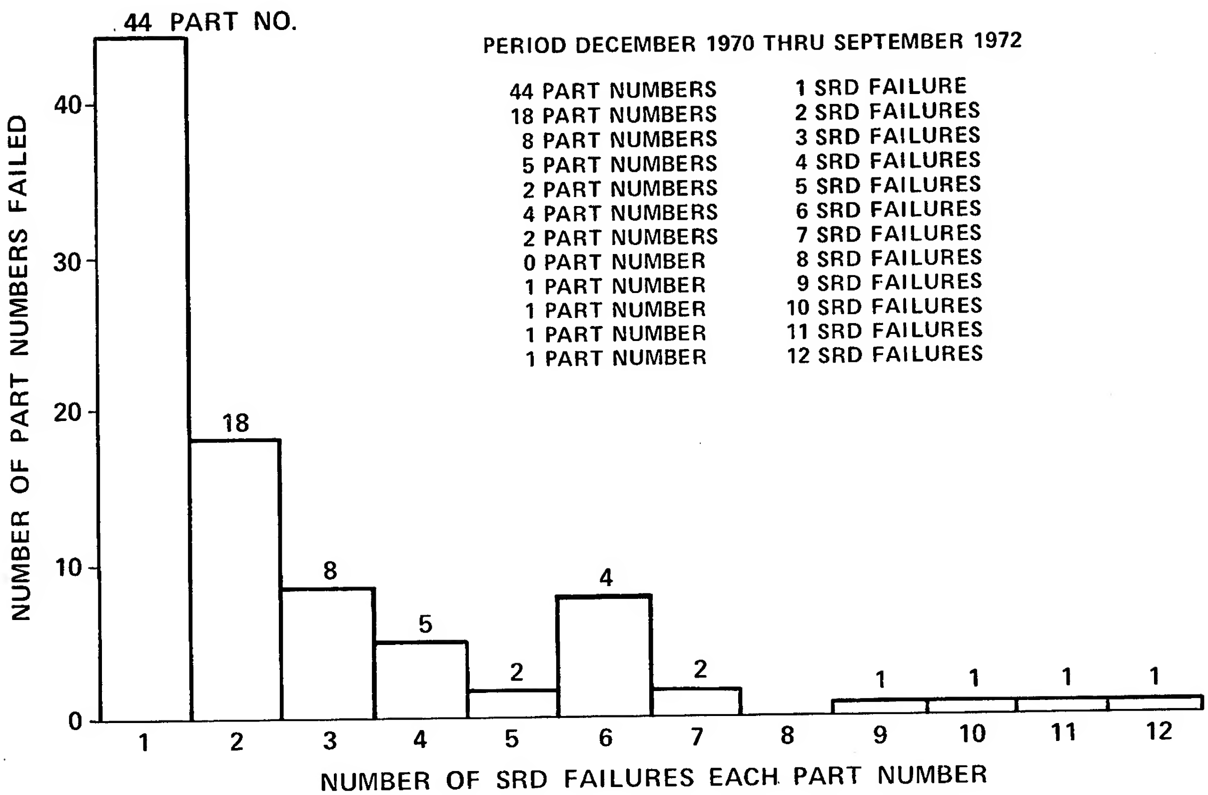


CHART 9 – REMOVAL FREQUENCY BY PART TYPE

	TRANSISTORS	MICROELECTRONIC DEVICES	CAPACITORS
PROCESS CONDITIONING (100%)	BURN-IN HIGH TEMPERATURE STORAGE TEMPERATURE CYCLING LEAK TEST ELECTRICAL SCREEN	BURN-IN CENTRIFUGE HIGH TEMPERATURE STORAGE TEMPERATURE CYCLING LEAK TEST ELECTRICAL SCREEN	ELECTRICAL TEST FOR ALL PARAMETERS ER SPECIFICATIONS USED WHERE AVAILABLE
LOT ACCEPTANCE TESTS GROUP A (SAMPLE)	ELECTRICAL CHARACTERISTICS AT ROOM, HIGH, AND LOW TEMP	ELECTRICAL CHARACTERISTICS AT ROOM, HIGH, AND LOW TEMP	ELECTRICAL TESTS
GROUP B (SAMPLE)	MECHANICAL ENVIRONMENTAL OPERATING LIFE STORAGE LIFE	MECHANICAL ENVIRONMENTAL OPERATING LIFE STORAGE LIFE	ENVIRONMENTAL TESTS LIFE TEST
IN-HOUSE TESTING (SAMPLE)	ELECTRICAL CHARACTERISTICS AT ROOM, HIGH, AND LOW TEMP TEMPERATURE CYCLING OPERATING LIFE	ELECTRICAL CHARACTERISTICS AT ROOM, HIGH, AND LOW TEMP TEMPERATURE CYCLING OPERATING LIFE	ELECTRICAL TESTS ENVIRONMENTAL TESTS

CHART 10 – TYPICAL QUALITY ASSURANCE REQUIREMENTS

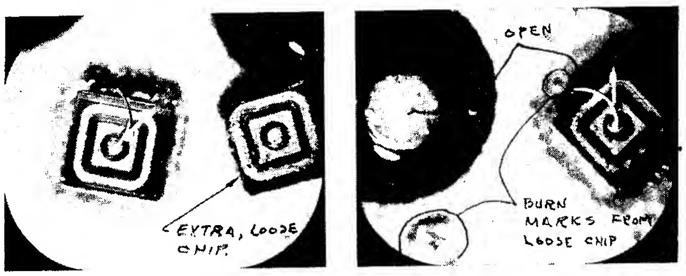


CHART 11 MICROPHOTOGRAPH OF FAILED TRANSISTOR

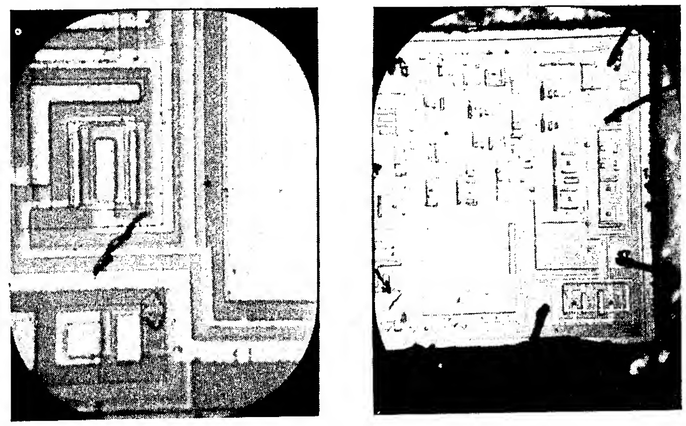


CHART 12 PARTICLE CONTAMINATION IN A MICROCIRCUIT

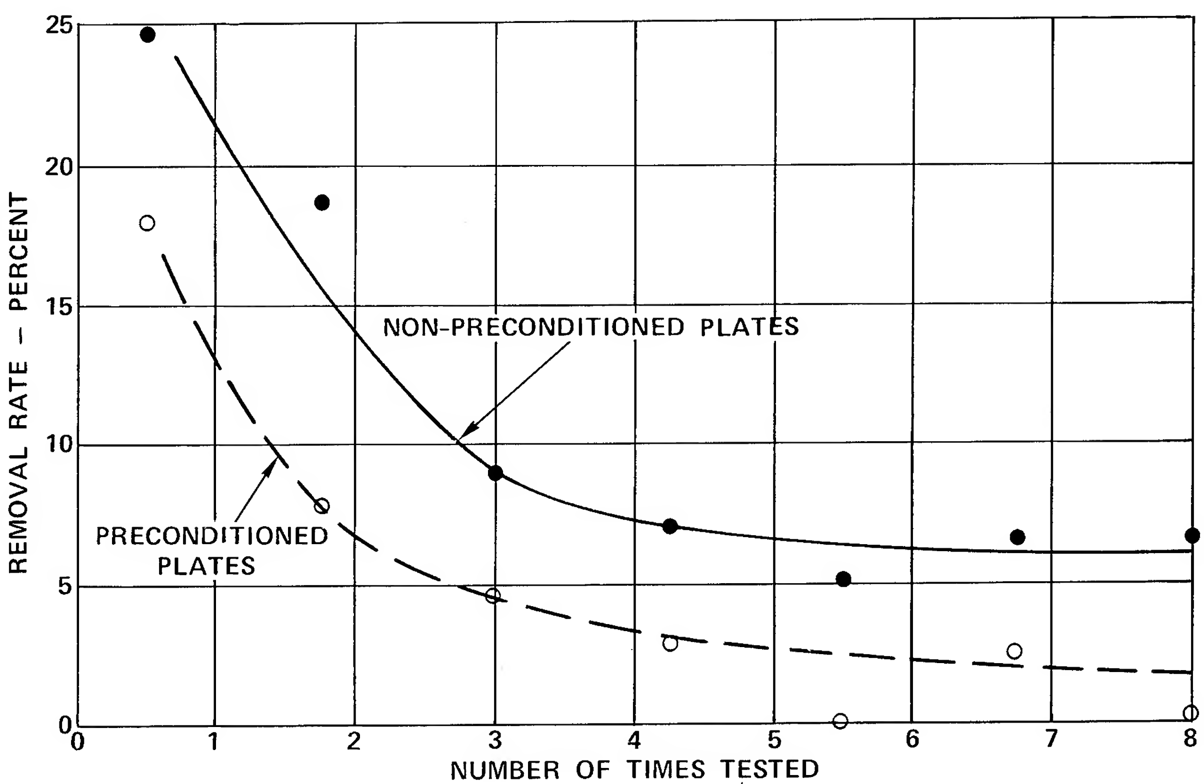


CHART 13 – AVERAGE ASSEMBLY REMOVAL RATES VERSUS NUMBER OF TIMES TESTED

QUESTION	YES	NO
IS BURN-IN A REQUIREMENT ON COMPONENTS?	9 – ALL PROGRAMS 3 – SELECTED PROGRAMS	2
IS BURN-IN CONDUCTED AT VENDOR?	11 – NORMALLY AT VENDOR	1 – RECEIVING INSPECTION (2 NOT APPLICABLE)
IS PRE-CAP VISUAL INSPECTION A REQUIREMENT?	4 – ALL PROGRAMS 3 – ALL PROGRAMS – SELECTED PARTS 3 – SELECTED PROGRAMS	4
IS CHANGE CONTROL INVOKED ON VENDOR PROCESSES?	1 – ALL PROGRAMS 1 – CAPTIVE LINE FOR MC'S 2 – PARTIAL	10
IS SOURCE INSPECTION/ CONTROL INVOKED?	10 – ALL PROGRAMS 1 – SELECTED BASIS ONLY	3
IS 100% AMBIENT TEST AT RECEIVING INSPECTION PERFORMED?	3 – ALL TYPES 2 – SELECTED TYPES ONLY – OTHERS SAMPLE TEST	9 – SAMPLE TEST ONLY
ARE ENVIRONMENTAL TESTS PERFORMED AT RECEIVING INSPECTION?	5 – SAMPLE BASIS 1 – SELECTED TYPES ONLY	8
IS INTERNAL EXAMINATION (DESTRUCTIVE) PERFORMED AT RECEIVING INSPECTION?	2 – ALL TYPES 1 – MC'S AND SELECTED SEMICONDUCTORS ONLY	10 1 – PERFORMED AT SOURCE

F90129 7102

CHART 14 – AEROSPACE MANUFACTURERS ELECTRONIC COMPONENT SEMINAR SUMMARY

• IS COMPONENT REJECTION RATE GETTING WORSE?	3 – YES 6 – NO CHANGE 5 – SLIGHT IMPROVEMENT
• IS USERS COMPONENT PROGRAM DOING THE REQUIRED JOB?	6 – YES 6 – YES, BUT COSTLY 2 – NO

CHART 15 – AEROSPACE MANUFACTURERS ELECTRONIC COMPONENTS SEMINAR SUMMARY

1. MIL-M-38510, LEVEL B, INVOKED ON ALL HYBRID MICROELECTRONIC DEVICES
2. PRE-CAP VISUAL TEST (MIL STD 883, METHOD 2072)
3. BOND STRENGTH TEST (MIL STD 883, METHOD 2011D)
   - INVOKED AS PART OF PRODUCTION PROCESS CONTROL CRITERIA
   - LARGE SAMPLE TEST PERFORMED IN HOUSE ON RECEIVED LOTS; LOT REJECTION IF SAMPLE TEST RESULTS IN FAILURE
4. LOOSE PARTICLE DETECTION
   - SAMPLE TEST AT SUPPLIER
   - LARGE SAMPLE TEST PERFORMED IN HOUSE ON RECEIVED LOTS; 100% SCREEN IF SAMPLE TEST RESULTS IN FAILURE
5. TEMPERATURE CYCLING – 100% SCREEN TO OPEN BONDS OF MARGINAL STRENGTH

CHART 16 – PART IMPROVEMENT PROGRAM

MORE RELIABILITY DATA FROM IN-FLIGHT SPACECRAFT

INDEX SERIAL NUMBER — 1071

E. E. Bean and C. E. Bloomquist
Planning Research Corporation
Los Angeles, California

and

Jay Finkelstein
Navy Space Systems Activity
El Segundo, California

Introduction

In a paper presented at the 1968 meeting of this symposium, selected results of a pioneering study devoted to the operational reliability of spacecraft were given (Reference 1). This study, upon which the 1968 paper was based, isolated, coded, and analyzed the reliability inherent in the operational records of 225 United States Spacecraft launched in 33 space programs prior to May 1966 (Reference 2).

As a part of an update of that earlier study, a complete revision to Reference 2 was published in November 1971 under the sponsorship of the Navy Space Systems Activity. Reference 3 is the updated report and contains data from 40 space programs on 304 spacecraft launched prior to January 1971. This paper is in the nature of a status report, briefly summarizing results to data which, for the most part, are contained in Reference 3. Other results obtained under NAVSPASYSACT sponsorship are contained in References 4, 5, and 6 and work is continuing in this area at the present time.

Background

The impetus for the first study was the lack of piece-part failure rates applicable to the space environment. Predictions of spacecraft on-orbit reliability based on data then available often gave results quite incommensurate with what could be observed from the large number of successfully orbiting spacecraft.

The initial study did a great deal to revise downward the previously accepted levels of part failure rates. By emphasizing actual on-orbit experience, and particularly the various incidents of anomalous behavior, it did a great deal more. Detailed descriptions including the occurrence time of 665 anomalous incidents were tabulated together with classification codes in eight categories. The need for data such as these was evidenced in the highly favorable response to the final study report.

Scope, Source, and Nature of Data

Major efforts were expended during the earlier study and the update to obtain basic, detailed data elements required for the analysis of spacecraft on-orbit reliability. In the earlier study the approach taken was to collect and analyze all reliability data from as many orbital spacecraft as possible within the cost and schedule constraints of the study. This approach tends to exclude spacecraft from highly classified programs for which the pertinent data are very nearly inaccessible. The update used essentially the same approach but further directed data collection activities to relatively complex unmanned spacecraft with intended missions of long duration. Overall, the data base covers approximately 40 percent of all U. S. spacecraft launches. The proportions covered yearly are shown in Figure 1.

The scope of the study and the update precluded the analysis of data at the detail level of raw telemetry reports or daily logs of the operational experience of spacecraft. Instead, the data search concentrated on obtaining summary reports and interviews from cognizant sponsoring agencies, contractors, and universities. As is to be expected in any large data-collection effort, the resultant documentation varied widely among programs and among launches of a specific program. Much of the needed documentation for the early programs (before 1962) either does not exist or was stored in archives where retrieval was not practical in the time available to the study.

Because of the wide variety of reporting formats encountered, a basic set of working papers was devised so that the available data for all launches could be compiled and reduced uniformly in a file called an engineering analysis report (EAR). General data elements recorded in an EAR for each launch included the mission description, launch vehicle, description of abortive launch (if any), launch date, orbit parameters, program objectives defined by the program office, and an overall evaluation of the in-flight performance. Reliability data elements included the spacecraft hardware breakdown to three levels of indenture (subsystem, equipment group/component, and piece parts); the number of powered hours, unpowered hours, or cycles experienced by the equipment for the three hardware levels; and a complete narrative description of anomalous behaviors, including the effect of the anomaly on the mission (catastrophic, negligible, modified by ground action, etc.), the effect on other hardware groupings, the implications on subsequent launches, and the assignable causes for the anomaly, if known.

In the generation of the EAR's, reliance on the referenced documentation was mandatory. A strong emphasis was placed on the recording of known values for all data elements, thus holding engineering assumptions to a minimum and thereby reducing potential biases from this source. This procedure does, however, reduce the total data sample somewhat and does not eliminate biases inherent in the source data. The major shortcoming in the data is that all anomalous incidents used in the analysis are "reported" incidents rather than the desired "occurring" incidents. There is considerable indirect evidence that for some spacecraft not all anomalous incidents were reported in the available documentation. This is much less of a problem in the update than in the original study.

Analysis of Anomalous Incidents

Selected results of the analysis of the two primary data elements (incidents of anomalous behavior and the operational profile of the spacecraft) are discussed in the following paragraphs for the updated data sample. Contrasts are drawn between the original study results and the update when they are instructive.

The total data base of anomalous incidents is profiled in Figure 2. Of the successfully launched spacecraft, 85 percent reported one or more anomalies whereas only 13 percent of the spacecraft reported 10 or more anomalies. The spacecraft added during the updating tended to have more reported anomalies than those analyzed earlier for three reasons: (1) the added spacecraft are generally more complex, (2) they are all long-term spacecraft, and (3) they are better documented.

Each anomalous incident recorded in the study contains the following information: (1) time of anomaly occurrence, (2) description of the incident, (3) incident cause, if known, (4) effect on the mission as a whole, (5) known corrective action taken on future flights or during the flight under consideration, and (6) clarifying remarks to place the incident in proper context.

To extract information from the narrative summary that would enable meaningful analysis, each anomalous incident was classified according to nine relevant characteristics. Four of these characteristics are selected for discussion here: (1) mission phase, (2) mission effect, (3) incident cause, and (4) spacecraft subsystem. The other five are identified and discussed in Reference 3.

Mission Phase

In general a spacecraft mission can be thought of as consisting of two major phases: launch and acquisition, and orbital or steady-state operation. The time intervals associated with these two phases are greatly different, with the launch and acquisition interval being very small relative to the nominal period of steady-state operation. Interestingly, of the reported anomalous incidents in the earlier sample, the numbers of such incidents are nearly equal for the two phases. In the combined sample about one-third of the anomalies occur in the launch phase and about two-thirds in the period of steady-state operation. These statistics reflect the large number of short term spacecraft in the original sample and the complete absence of them in the update. Of those anomalies added in the update nearly 90 percent were from the orbital or steady-state phases.

The frequency distribution of anomaly occurrence times as shown in Figure 3 indicates the extreme importance of the early portion of a spacecraft mission to its ultimate reliability.

Mission Effect

The effect of most anomalies on the spacecraft mission is very small. To classify the anomalous incidents according to their effect on the spacecraft mission, five categories were defined based on a judgment of the effect of each incident on the overall mission, had it occurred in isolation. On a zero to one scale, where one represents catastrophic failure of the mission and zero represents no effect, the five categories or groups may be defined as follows: Group 1 (0 to ϵ); Group 2 (ϵ to 1/3); Group 3 (1/3 to 2/3); Group 4 (2/3 to $1-\epsilon$); Group 5 ($1-\epsilon$ to 1). Of the 1190 sample incidents, only two could not be assigned a severity classification in this manner. The percentages of the observed anomalies according to the five severity classifications are presented in Figure 4. As can be seen from the exhibit, more than 50 percent of the reported anomalies had little or no effect (severity group 1) on the accomplishment of the spacecraft mission. The distribution of the anomalies among the five mission effect categories is virtually unchanged between the original and updated samples. It should be noted that very few spacecraft lifetimes end as a result of catastrophic failures(severity group 5). Spacecraft lifetime is far more likely to be determined by the cumulative effect of lower severity anomalies.

Anomaly Cause

Each recorded anomaly was investigated to determine if its cause was assignable, nonassignable, or unknown. An assignable cause was attributed to a specific anomaly if that incident could have been prevented by taking some action well within the state of the art prior to launch, or if it was the direct result of some other anomalous behavior. If such was not the case, the incident was classified as nonassignable. The unknown category contains those incidents wherein insufficient information was available to make a judgment. Figure 5 shows the percentages of the 1190 anomalies failing in each of these three broad categories. This distribution again is virtually identical for the original and updated samples.

The assignable cause group is of significant interest because, to reduce the number of anomalous incidents on spacecraft and thus improve reliability, it is necessary to remove the cause of the anomalous behavior. The other two groups offer little in the way of improving spacecraft postlaunch reliability, either from lack of data or from lack of any evident corrective action. Thus, those incidents from which an assignable cause is evident are worthy of a more detailed examination in an effort to discover the contribution they could make in pointing out correctable trouble areas.

The assignable cause category may be considered to be composed of six general areas:

1. Design. Included in this area are RFI and sensitivity problems, unanticipated wearout, or degradation as a result of time or known environmental conditions. The anomalies may be electrical, mechanical, thermal, or system-related.

2. Manufacture. Included in this area are such causes as faulty parts or materials, contamination, faulty solder joints or other connections, quality control, etc.

3. Operation. Incidents included in this area are the result of human error in the spacecraft control function, usually in commanding, programming, or calibrating the spacecraft.

4. Another Anomaly. Included in this area are those anomalies that occurred as the direct result of some previous anomaly.

5. Nonanomalous Behavior. Some incidents included in the sample are reported mainly for interest and cannot, in the strict sense, be called anomalous behavior. These include incidents such as

failure of equipment operating beyond their intended lifetime and miscellaneous equipment opportunely launched as part of development testing.

6. Acts of God. This area does not represent anomalous behavior per se, but reflects spacecraft behavior that is the result of unanticipated external sources, e.g., meteoroid bombardment.

An intensive survey of the 415 reported anomalous incidents with evident assignable causes, together with the foregoing considerations, leads to the detailed breakdown of assignable causes shown in Figure 6. The six primary categories are shown under "all assignable causes," together with the number of anomalous incidents classified as belonging to that category. The two primary categories of "design" and "manufacture" are further subdivided. Each category and subcategory contains both the total number of such incidents in the sample and the percentage of such incidents.

The various subcategories under "design" are indicative of certain reported assignable causes as follows: (1) the subcategory "RFI, etc." includes all anomalous incidents attributed to inadequate RFI design, noise sensitivity, spurious commands and transients; (2) the three subcategories "system, mechanical, thermal" include incidents arising from inadequate design in the spacecraft/environment or subsystem interfaces, in deployment or structural integrity, and for proper spacecraft thermal balance (usually reported as overheating problems); (3) the category "electrical component" refers to anomalies attributed to inadequate design of a receiver, encoder, horizon sensor, etc.; (4) "unanticipated wearout or degradation" is attributed to anomalies where, for example, a battery wears out before anticipated, or where other components or parts do not have the capability to survive either the normal environment or specified time; and (5) the three remaining subcategories ("radiation," "launch vibration and shock," and "atmospheric conditions") indicate that the anomaly resulted from a design inadequate to withstand these conditions. The various subcategories are exhaustive and mutually exclusive; hence, each "design anomaly" is attributed to one and only one of the subcategories, depending on which seemed most nearly appropriate.

The subcategories under "manufacture" are intended to be somewhat more explicit in the reported assignable cause. Included under "fabrication, Q.C., etc." are anomalies like cold or loose solder joints, loose connections, missing parts, and defects. "Contamination" covers the relatively high occurrence of clogged hydraulic lines, excess moisture, foreign matter in valves, and the like. "Faulty parts or materials" indicate such items as foreign matter in a transistor or use of degraded propellants.

Figure 6 indicates that the most common cause of spacecraft anomalies is inadequate design, representing well over 60 percent of all incidents having assignable causes. Manufacturing problems accounted for 20 percent and spacecraft operation for 10 percent of all incidents with assignable causes. The remaining anomalies (10 percent) were distributed among secondary failures, anticipated anomalies, and acts of God.

Spacecraft Subsystem

To identify the most troublesome functional aspect of spacecraft, the anomalous incidents were classified according to a number of functions that could in turn be classified according to spacecraft subsystems.

The five spacecraft functions with the highest incidence of anomalies per function are as follows: (1) technological payloads, (2) data point sensing and monitoring, (3) data storage, (4) life support, and (5) active thermal control. The appearance of (1), (2), and (4) in the list is not unexpected, as these functions are generally monitored and controlled most closely. Data storage and active thermal control may be more indicative of true problem areas.

This ranking is stable from the original sample to the updated sample. The only differences are that in the original listing the ranks of the first two functions were interchanged and that orientation sensing held fifth rank rather than active thermal control as is the case here.

The spacecraft functions with the lowest incidence of reported anomalies per function are (1) basic structure, (2) spacecraft separation, (3) power distribution, (4) command decoding, and (5) navigation. The difference between this list and the one based on the original sample is that the command receiving function was included and ranked number 2.

The telemetry and data handling subsystem accounts for over one-third of all reported incidents; the timing, control, and command subsystem, the power supply subsystem, the attitude control and stabilization subsystem, and the payload subsystem account for approximately 14 percent of the anomalies each. The remaining 11 percent of the anomalies are distributed among propulsion, environmental control, structure, and unknown subsystems. This distribution of anomalies is not substantially different from that maintaining in the original sample.

Examination of mission effect in conjunction with the subsystem category indicates that anomalous incidents occurring on the structure and power supply subsystem are more likely to be seriously degrading to the mission. Environmental control and telemetry and data handling subsystems are relatively less likely to affect mission accomplishment adversely.

Hardware Element Reliabilities

The reliability of a hardware element considered herein is the probability of survival. The probabilities were derived for three tiers of spacecraft hardware elements; subsystems, components, and piece parts. The probabilistic nature of the reliability investigation necessitates the making of numerous assumptions and some selection of the available data to arrive at meaningful results. The assumptions and data selection are discussed in context below.

Two probabilities of interest were computed: first, the probability of failure during launch, and second, the probability of hardware element survival for t hours of orbital operation. Under the assumption that each identically named hardware element has

an equal probability of failure during launch, irrespective of mission, then q , the probability of a hardware element failure during launch, is estimated by

$$\tilde{q} = \frac{\ell}{N} \qquad (1)$$

where ℓ = Number of hardware element failures during launch

N = Total number of hardware elements in the sample

The probability of hardware element survival during orbital operation, R(t) , is computed under the major assumption that the time to failure is adequately described by an exponential distribution: that is,

$$R(t) = \exp(-\lambda t) \qquad (2)$$

where λ = the hardware element failure rate.

This assumption is widely used in reliability literature and practice, especially for most electronic hardware elements found in spacecraft. The data generated in the referenced study did not support the use of alternative assumptions.

In instances when Equation (2) applies, it is well known that the best estimate of λ for a particular hardware element type is given by

$$\tilde{\lambda} = \frac{f}{\sum_{i=1}^{n} t_i} \qquad (3)$$

where n = number of equivalent hardware elements under observation

t_i = survival time of the ith such element

f = total number of failures observed

The formulations for determining confidence intervals for $\tilde{q}$ and $\tilde{\lambda}$ are well known and will not be repeated here.

A good deal of effort was expended to obtain survival hours for a great variety of hardware elements, particularly at the component and piece-part level. The key step in this process was determining and listing components at a level sufficiently high so that their operational history could be readily determined and yet sufficiently low so that it was reasonable to assume that their normal operation would be precluded on occurrence of a piece-part failure. The components and piece-part survival hours, together with any known failures, were portions of the basic data sheets (EAR's) generated during the study. By integrating over the component operating histories within a subsystem, the subsystem operating history was determined. By deduction, the operating histories of piece parts within the component were determined.

A failure was attributed to a piece part if, and only if, it was known to have failed in a catastrophic manner for no evident cause. Failures were attributed to a component in the same manner, essentially by treating the entire component as a big piece part.

The primary ground rule throughout the calculations was to use known values only. For example, within the sample there were 497 transmitters for which operational histories were complete. Cumulatively they survived at least 3,210,000 hours and exhibited no launch failures and 11 orbital failures. In addition, it was known that each of these figures is in fact higher than those presented but, because of inadequate data, it is not known by how much. Therefore, some caution is urged in interpreting the resultant estimates of q and λ .

Subsystems

Figure 7 presents the best estimates and confidence limits for the launch failure probabilities and in-orbit failure rates for spacecraft subsystems. The subsystem list presented in the exhibit is an expedient used to avoid listing recognizable subsystems (i. e., those traceable to a specific program). Nevertheless, for large system planning considerations, it seems helpful to have some indication of gross, average launch failure probabilities and in-orbit failure rates for spacecraft subsystems. A subsystem failure is defined as some anomalous behavior associated with the subsystem, the result of which is to reduce mission effectiveness by at least two-thirds of its potential effectiveness.

The parameters are felt to be reasonably indicative of failure propensities (or conversely, reliability) of spacecraft subsystems. The most important bias underlying this analysis results from the tendency to report details of a subsystem's operation if it exhibits anomalous behavior and not to include such details if its operation is essentially perfect. This situation tends to raise the parameter values shown in Figure 7. The bias is particularly noticeable in the environmental control subsystem, where sufficient information was available for only 28 of these subsystems. Although no failures are shown, a number of incidents with little or no effect were noted. The bias is felt to be minimal in the other subsystems because of their degree of criticality to any degree of mission success.

A comparison of the updated sample with the original sample indicates an almost unanimous reduction in subsystem failure rates and probabilities of launch failure. Reductions in subsystem failure rates extend over all subsystems and range from 23 to 84 percent with the average being 55 percent. The payload subsystem exhibits a higher probability of launch failure in the updated sample than in the original one; for all other subsystems the updated sample indicates lower launch failure probabilities.

Components

Figure 8 provides estimates of the launch failure probabilities for those components in the updated sample with one or more launch failures. Only 14 component failures were observed in the launch phase in the entire updated sample. Transponders with three failures, and sequencers and receivers with two failures each are the only component types with more than one launch phase failure. The only observed launch phase failures in the original samples were the two associated with the receivers. This large augmentation in the data base probably reflects better reporting procedures rather than declining reliability.

Figure 9 presents the estimates of components on-orbit failure rates, and 90-percent confidence intervals, for all components in the updated sample with one or more failure and 3,000 or more survival hours on orbit.

The majority of components considered exhibited no failures either during launch or in orbital operation. The most failure-prone component in both the updated and original samples appears to be the magnetic tape unit. The updated sample indicates 38 failures occurring on 132 units observed. The current failure rate of 40 failures per million hours is significantly higher than the 28 failures per million hours observed in the original sample. DC/DC converters were the only other component exhibiting a higher failure rate in the updated sample than in the original sample.

Piece Parts

At the piece-part level, many more assumptions are required with respect to operating hours, because telemetry data are insufficient to describe the operational history of many specific piece parts. The assumption used in the analysis is that, so long as a component is completely operable, so is every piece part. When a component is removed from the sample because of anomalous behavior, the piece parts were removed if there is any suspicion that the anomaly is caused by a piece part. The result is that the operating hours for piece parts represent minimum part hours within the limits of the input data.

As before, a failure is entered in the calculations only if the part failed catastrophically for no evident cause. Of the 740 anomalous incidents in which a determination could be made as to part-responsibility, only 23 percent represented part failures, 21 percent were non-catastrophic part failures and 56 percent were not part related. The number of part failures is probably lower than the true value for the following reasons: (1) some part failures are never detected because of minimal effect, low-level redundancy, etc.; (2) some detected part failures are not reported--an inevitable situation where no formal procedure exists for such reporting; (3) some anomalies strongly suspected as originating from a part failure simply cannot be isolated to the particular part; and (4) many anomalous behaviors are noted for which it is unknown whether or not a piece-part failure is involved. It is a fact, however, that the updated sample, which in general is better documented, indicates a significantly lower proportion catastrophic part failures and a significantly higher proportion of anomalies which are definitely not part related.

No table for probability of failure during launch is given as only one capacitor, one transistor, one transducer, and one relay were observed to have failed during this phase. The transducer and relay failures were added in the update. As in the earlier study, no statistics are reported for squibs, cartridges, and other essentially one-shot devices. Determination of actuation and exact redundancy configurations is simply not possible; however, no anomalous behavior on any spacecraft studied could be attributed to these devices.

Figure 10 presents the estimates of piece-part failure rates and 90-percent confidence intervals for those parts which exhibited one or more failures in 3,000 or more survival hours of orbital operation. When these results are compared with the results given in the earlier paper, then (with only one exception, thermistors) the estimated piece-part failure rates are lower in the updated sample than in the original sample and are usually lower by a substantial amount. The four higher population discrete piece-parts illustrated in Figure 11 indicate the trend.

Figure 12 compares some of the piece-part failure rates given in Figure 10 with rates commonly used in reliability assessment calculations. The commonly used rates are derived from four sources: (1) rates used by reliability analysts at Planning Research Corporation in assessment activities (Reference 7); (2) the Earles and Eddins failure rate tabulations (Reference 8); (3) Mil-Handbook 217/A (Reference 9); and (4) Minuteman failure rates (Section 7 of Reference 9). All rates are either generic failure rates (i.e., no application K-factor is applied) or rates purported to be applicable in the space environment. Minimum values were selected in all cases; the minimum and maximum rates presented in the exhibit, therefore, are with respect to the previously-named sources. Where a reasonably comparable minimum failure rate from Mil-Handbook 217/A is available, it is also tabulated because it is the most widely used reference work for failure rates. High-population parts generally have a much lower failure rate than Mil-Handbook 217/A and, except for Minuteman parts, lower than all in common use. None of the failure rates estimated from on-orbit data are higher than the upper end of the interval defined in Figure 12. This is at least in part a result of the biases mentioned previously. It seems apparent, however, that high-population parts (capacitors, diodes, resistors, and transistors) have failure rates considerably lower than those generally assumed appropriate for space application. Furthermore, the failure rate reduction factors tabulated above indicate that the estimates of Figure 10 could be expected to decline even further since survival hours are accumulating faster than failures for virtually all part types.

On-Off Cycling and Dormancy

Although provision was made for collecting data pertinent to on/off cycling and dormancy in the original study the data were simply too sparse to provide results or to even attempt an analysis. In the update, however, particular emphasis was placed on securing data pertinent to this question and in analyzing the data that were collected. Unfortunately, the analytical results were not clear-cut.

On-Off Cycling

Defining the subject matter in clear and unambiguous terms is the most difficult part of the problem. This difficulty is a function of the dynamic behavior of nearly all orbiting spacecraft and particularly the more recent and complex satellites. Each major subsystem may be characterized by a number of operational modes, many components are normally subject to cyclical operation (for example, the record and playback cycle of tape recorders, battery charge and discharge cycles, etc.) and configuration changes via the ground/spacecraft link are common on nearly every pass. To compound the problem there are rarely sufficient data to quantify any of the parameters associated with the above operation (time spent in playback mode or record modes, number of playbacks, operational hours per mode, etc.).

The approach taken to surmount this difficulty is, again, to place reliance on "known" values and to

keep engineering assumptions to an absolute minimum. When available program documentation provides clear and reasonably straightforward data regarding the cycling of spacecraft components, it is utilized; otherwise it is not.

Cycling data were found for nearly 200 components. These data include (1) the component type, (2) the number of parts in the component--discrete or integrated circuit, (3) total on-orbit survival time, (4) power-on time, (5) number of cycles, and (6) number of anomalies.

The component type is quite variable ranging from a 20-piece-part power convertor to an entire spacecraft consisting of some 20,000 electronic piece-parts. Survival times ranged up to 25,000 hours with power on time ranging from practically zero percent to practically 100-percent of total survival time. The number of cycles varied from one to 8500. The most common number of anomalies per component was zero but was actually 30 for one of the components which represented an entire spacecraft.

The survival hours represent the time that the component under consideration was known to be operable. Power-on time is the number of hours that full, nominal power was applied to the component. Survival hours minus power-on hours gives the time that the component was dormant or on inactive standby.[1] The number of cycles is essentially the number of turn-ons, i. e., switching from inactive standby to full, nominal power. It is not too unreasonable to assume that the on periods in each cycle are approximately equal.

The most notable feature of these data, taken as a whole, is the general lack of anomalistic behavior associated with the cycled components and the fact that none of the recorded anomalies can be attributed, unambiguously, to the cycling itself or to the dormant period of the component's operational profile.

Comparing the on/off cycling data to the survival data including all kinds of operation there is no striking or statistically significant difference. There are, for example, 51 transmitters represented in the on/off cycling data with a total of 455,779 survival hours, no catastrophic failures,[2] and 27,517 on/off cycles. In terms of survival hours this represents a 90-percent confidence interval on the failure rate of 0 to 5.1×10^{-6} failures per hour compared to the interval of 1.0 to 4.7×10^{-6} failures per hour found for all transmitters. These results are not unexpected given that the two populations are essentially equal in terms of failure rate. To deduce from this example that cycled and uncycled components, which are otherwise similar, have the same failure rates is not warranted, however, on two counts. First, it is not unlikely that all the transmitters included in the analysis were cycled to some extent, those represented in the on/off cycling data being simply the transmitters for which quantitative cycled data are available. The second problem is the sparsity of failure data which tends to make all failure rate comparisons somewhat nebulous.

1 The terms "dormant" and "inactive standby" are considered to be synonymous in this report.

2 Although there are 10 anomalies recorded against six integrated circuit transmitters none of these resulted in the termination of transmitter operations.

Thus, although no clear pattern emerges from the data which could be used to reject the hypothesis of equal component failure rates under cycling and steady state operation, equality is not therefore demonstrated. No general decision on the impact of cycling can be reached either way at this time on the basis of currently available data. It is rather clear, however, that cycled components in general do not have "order of magnitude" worse failure rates than their non-cycled counterparts.

There may well be compensating tendencies in the cyclic mode of operation in that turning a component on and off may be detrimental to reliability whereas periods of no (or reduced) stress may be beneficial. The detrimental effect of on/off switching was found in the analysis of Reference 5 for the various scientific experiment packages of an observatory class satellite; the beneficial effects of dormancy were not. The evidence from the data of this study does indicate that a cycling rate in excess of 0.1 cycles per hour is worse, in terms of reliability, than cycling less often.

To conclude, it is not clear on the basis of the empirical data of Reference 3 whether cycling per se is detrimental to spacecraft components, compared to steady state operation; it is reasonably clear, however, that if spacecraft components are to be cycled it is desirable to reduce the cycling rate.

Dormancy

As indicated earlier no components or piece parts are known to have failed when they were in a dormant condition or on standby. An explicit calculation of dormant failure rates is therefore not possible. The numbers of hours accumulated against some items, however, indicate that a rather low rate would be appropriate.

A tabulation of the upper 90-percent confidence limit on the dormant failure rate for selected components and piece parts is given in Reference 3. The basic data and method of calculation are as outlined previously for average on-orbit failure rates and indicate an upper 90-percent confidence limit generally higher than that found for the overall on-orbit failure rates. The generally higher dormant failure rate limit simply reflects the reduced amount of data available. For some components and piece parts, however, the failure rate limits are quite comparable. For six hardware elements the dormant failure rate limit is actually less than the overall on-orbit limit. These six elements are: DC/DC Converters, Magnetic Tape Units, Transmitters, Transponders, Vidicon Cameras and Switches.

Figure 13 gives the failure rate statistics for these six elements. For DC/DC Converters, Transponders, and Switches the upper failure rate confidence limits are about equal which only indicates that dormancy is probably no worse than general on-orbit experience. Vidicon cameras appear to profit from dormancy since the upper limit on dormant failure rate is less than the expected value from general on-orbit experience. The Magnetic Tape Units and the Transmitters, however, indicate a clear cut failure rate reduction from dormant operation; a factor of nearly 10 to one is indicated for the Magnetic Tape Units and of better than three to one for the transmitters. It is therefore reasonably clear, and made clear by demonstration from actual field data, that dormant failure rates are lower for some components than general

on-orbit rates and hence lower than operating failure rates.[1] Whether additional data could extend this conclusion to other components cannot be reasonably conjectured at this time.

Other Observations of Interest

There were 37 incidents of degraded or intermittent piece-part operation reported in the updated sample for which no assignable cause was evident. Considering that the combined sample contains 141 catastrophic piece-part failures, this would imply that if a piece-part misbehaves, the probability is on the order of 1/5 that it will not be a random catastrophic failure.

The apparent self-healing capability of spacecraft commented on in the earlier study is still present in the sample of this study. There are in the total sample 38 instances of anomalous behavior, involving 27 different spacecraft that were reported to have been completely recovered at a later date. Recovery times vary from a few milliseconds to more than 5 months. As indicated in Reference 1 the only unifying characteristic of these anomalies seems to be their electronic nature.

In the sample of this study redundancy played an important part in reducing the effects of an anomaly. There are 48 incidents where simple redundancy prevented a more serious effect. In 40 other incidents the seriousness of the anomaly was alleviated by "backup" other than redundancy, either an alternate means of achieving the same function or "work-around" procedures developed by ground control. In the original study these numbers were 25 and 15, respectively.

A final observation is that wearout of hardware units is not a significant problem. Among the anomalous incidents of this study, only five such incidents were noted. Two involved batteries, two were special purpose relays, and one was a solar X-ray detector.

Conclusions

The classification of anomalous incidents reported on the successfully launched spacecraft (87 percent of all spacecraft in the updated sample) result in the following major conclusions.

1. Eighty-seven percent of the successfully launched spacecraft reported one or more incidents of anomalous behavior; 12 percent reported 10 or more such incidents.

2. Seventy-one percent of the anomalies are reported in the orbital or steady-state phase of the spacecraft mission.

3. Eighty-nine percent of the reported anomalies have little or no effect on accomplishment of the spacecraft mission.

4. Two subsystems account for approximately one-half of the reported anomalies. The telemetry and data handling subsystem accounts for 34 percent of the reported anomalies and the payload subsystem accounts for 18 percent. Forty percent of the anomalous incidents are distributed essentially equally between timing and control, power supply, attitude control and stabilization and the remaining 10 percent are also distributed essentially equally among the propulsion, environmental control, structure, and unknown subsystems.

5. Over three-fourths of the anomalous incidents reported are electrical in nature as opposed to mechanical, chemical, unknown, etc. Only 15 percent of the incidents are catastrophic part failures; 13 percent are noncatastrophic part failures (degraded, intermittent, etc.); 35 percent are non-part related; for the remainder, no determination could be made as to whether a part is involved or not.

6. Fifteen percent of the incidents occurred for no apparent reason; 35 percent were the result of an assignable cause. For the remaining incidents no conclusions could be drawn as to the assignability or nonassignability of cause of failures. For those incidents having assignable causes, nearly 65 percent were attributed to various aspects of the spacecraft design, 14 percent to manufacture, and 9 percent to spacecraft operation; the remaining 12 percent were distributed among secondary failures, anticipated "anomalies," and acts of God.

Estimates of the spacecraft element reliability parameters, failure rate and probability of failure, in addition to their tabulation as given at the end of the paper, result in the following general conclusions.

1. The updated sample indicates that the power and attitude control and stabilization subsystems have the highest in-orbit failure rate among the subsystems. The propulsion, environmental control, and structure subsystems have no reported anomalies during orbit. Except for the telemetry and data handling and environmental control subsystems (neither have any reported anomalies during launch), the basic spacecraft subsystems exhibit essentially equal probabilities of failure during launch.

2. The majority of the components considered in both samples exhibited no failures either during launch or in orbital operation. The most failure-prone component appears, as it did in the earlier study, to be the magnetic tape unit with 38 failures occurring on 132 units observed. The failure rate for magnetic tape units in the combined sample is 40 failures per million hours, a significant increase over that reported in the earlier sample (28 failures per million hours). Only one other component had an increased rate compared to the rate reported earlier.

3. There are only four failures attributed to piece parts during launch (one each of capacitors, relays, transducers, and transistors) and only 40 during orbital operations. Forty-two part types are included in the study. High population parts (capacitors, diodes, resistors, and transistors) have significantly lower on-orbit failure rates when compared to those reported in the original study. The on-orbit failure rates of capacitors (0.87 per billion part hours) diodes (1.2 per billion part hours), resistors (0.21 per billion part hours), and transistors (0.65 per billion part hours) reflect the large number of observed units and operating time and the relatively few observed on-orbit failures.

The analysis of on/off cycling gives no clear evidence of a supposed detrimental effect on reliability of cycling spacecraft components as opposed

1 This is true since the on-orbit rates are based on a combination of powered and unpowered hours in unknown ratios.

to a steady state operation. The data indicate, however, that for cycled components a rapid cycling rate is more adverse than a slower one.

The effect of dormancy on reliability is also ambiguous. The analysis of this factor does demonstrate conclusively, on the basis of empirical data, that magnetic tape units and transmitters have a much higher operating failure rate than dormant failure rate. No failures or anomalies were identified which could be attributed to dormancy.

REFERENCES

1. Bean, E. E. and C. E. Bloomquist, "Reliability Data From In-Flight Spacecraft" Proceedings of the Annual Symposium on Reliability, January 1968.

2. Planning Research Corporation, PRC R-948, Study of Reliability Data From In-Flight Spacecraft, (NASA CR 84628) March 1967.

3. Planning Research Corporation, PRC R-1453, Reliability Data From In-Flight Spacecraft; 1958-1970, (AD 889943L), 30 November 1971.

4. Planning Research Corporation, PRC D-1799, Reliability Analysis and Synthesis, Status Report, 30 October 1970.

5. Planning Research Corporation, PRC R-1435, The Effects of Ground Storage, Space Dormancy, Standby Operation, and On/Off Cycling on Satellite Electronics, 28 May 1970.

6. Planning Research Corporation, PRC R-1293, An Investigation of On/Off Effects on Equipment Operating In Space, October 1969.

7. Planning Research Corporation, PRC R-919, Seventh Reliability Assessment for Orbiting Geophysical Observatories, 15 November 1966.

8. Earles, D. R., and M. F. Eddins, Failure Rates, Avco, April 1962.

9. Military Standardization Handbook, Mil-Handbook 217/A, Reliability Stress and Failure Rate Data For Electronic Equipment, 1 December 1965.

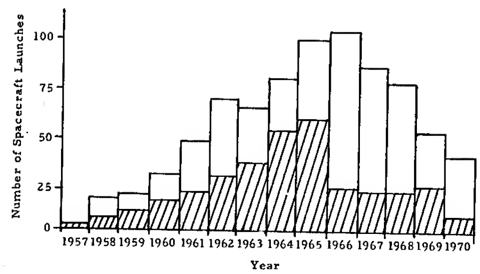


FIGURE 1 - DISTRIBUTION OF ALL U. S. SPACECRAFT LAUNCHES, AND THOSE IN THE SPACE DATA BANK, THROUGH 1970

FIGURE 2 - TOTAL DATA BASE OF ANOMALOUS INCIDENTS

	Updated Sample	Original Sample
Number of Spacecraft	304	225
Unsuccessful Launches	40	27
Spacecraft with no Reported Anomalies	40	34
Spacecraft with Reported Anomalies	224	164
Number of Anomalies Reported	1,190	665

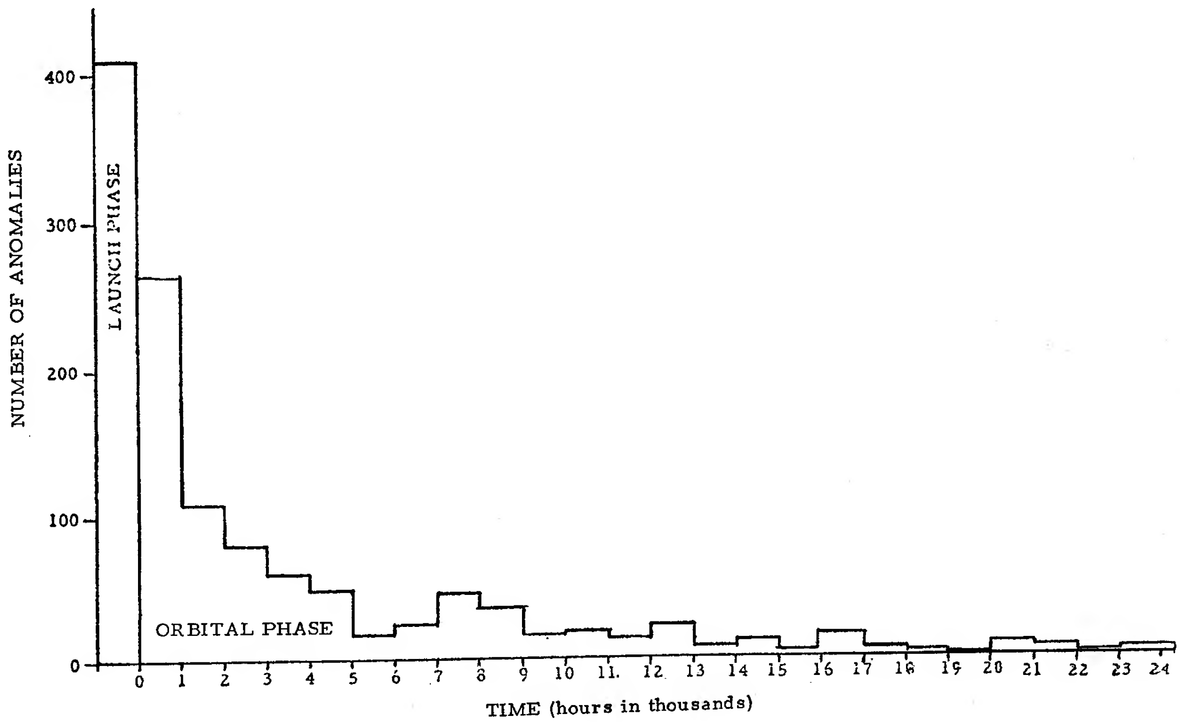


FIGURE 3 - FREQUENCY DISTRIBUTION OF ANOMALY OCCURRENCE VERSUS TIME

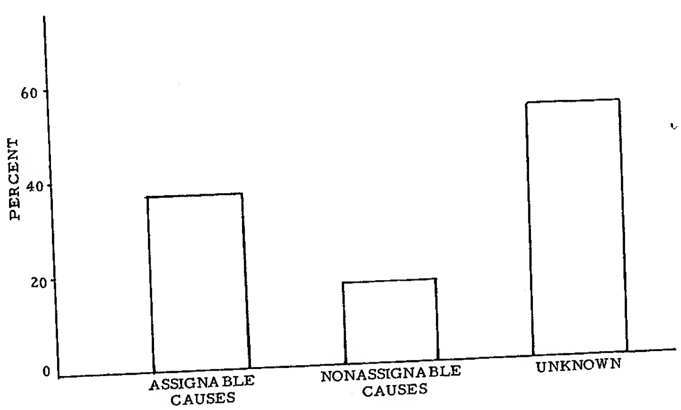


FIGURE 5 - ANOMALIES CLASSIFIED BY INCIDENT CAUSE

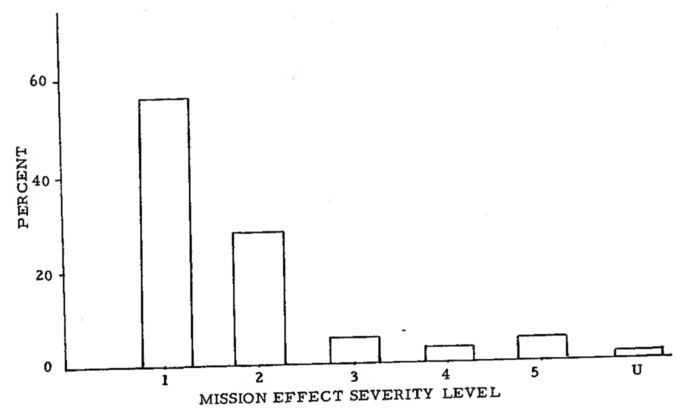


FIGURE 4 - ANOMALIES CLASSIFIED BY MISSION EFFECT

FIGURE 6 - DETAILED BREAKDOWN OF ANOMALOUS INCIDENTS BY ASSIGNABLE CAUSE AND MISSION TERM FOR SUCCESSFULLY LAUNCHED SPACECRAFT OF THE COMBINED SAMPLE

Causes	Numbers in Sample	
All Assignable Causes	415	
Design	269	
RFI, etc.		(67)
System		(49)
Electrical Component		(51)
Mechanical		(23)
Thermal		(28)
Unanticipated Wearout or Degradation		(13)
Radiation		(14)
Launch Vibration and Shock		(11)
Atmospheric Conditions		(13)
Manufacture	58	
Fabrication, Q.C., etc.		(29)
Contamination		(17)
Faulty Parts or Materials		(12)
Operation	38	
Another Anomaly	16	
Nonanomaly	30	
Acts of God	4	

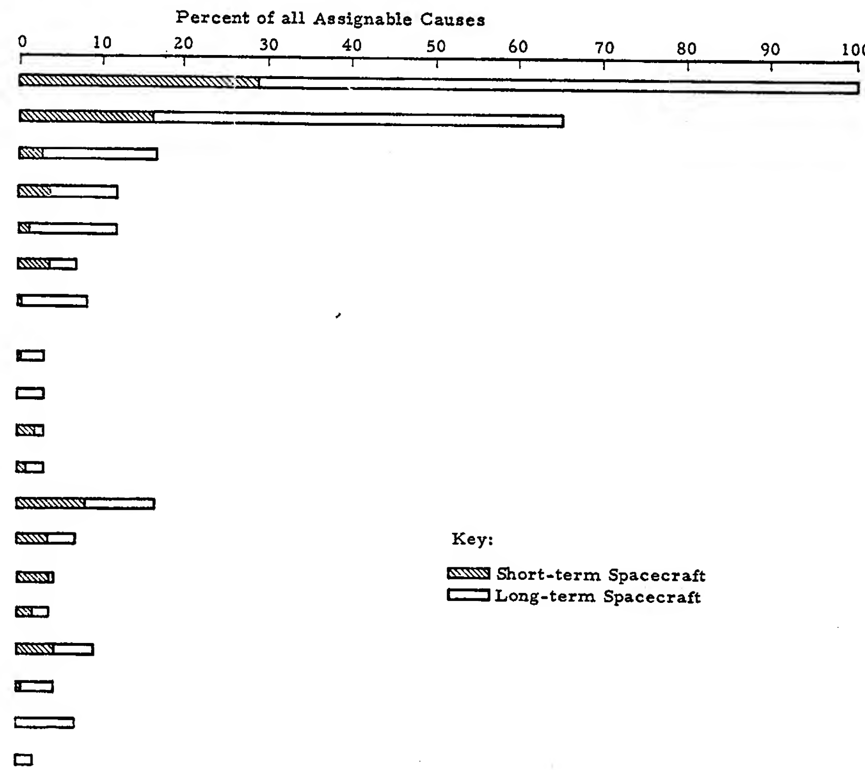


FIGURE 7 - SPACECRAFT SUBYSTEM RELIABILITY PARAMETER ESTIMATES AND 90-PERCENT CONFIDENCE INTERVALS FOR COMBINED DATA SAMPLE

Spacecraft Subsystems	Probability of Failure During Launch			In-Orbit Failure Rate (Failures/Million Hours)		
	q_1	$\hat{q}$	q_2	λ_1	$\hat{\lambda}$	λ_2
Timing, Control, and Command	0.0062	0.018	0.035	0.89	2.7	5.9
Telemetry and Data Handling	0	-	0.019	0.22	1.2	3.9
Power	0.0058	0.017	0.032	4.9	8.1	13
Attitude Control and Stabilization	0.0069	0.020	0.039	1.3	3.8	8.8
Propulsion	0.0072	0.027	0.057			
Environmental Control	0	-	0.079	0	-	5.4
Structure	0.0058	0.018	0.032	0	-	1.3
Payload	0.0044	0.013	0.025	0.033	0.64	3.0

$P\left\{q_1 \le q \le q_2\right\} = 0.90$ $\qquad$ $P\left\{\lambda_1 \le \lambda \le \lambda_2\right\} = 0.90$

Note: The subscript 1 denotes the lower confidence limit, the subscript 2 denotes the upper confidence limit, and the caret denotes the mean value.

FIGURE 8 - PROBABILITY OF FAILURE DURING LAUNCH AND 90-PERCENT CONFIDENCE INTERVALS FOR SELECTED SPACECRAFT COMPONENTS FOR THE COMBINED DATA BASE

Component	Probability of Failure During Launch		
	q_1	$\hat{q}$	q_2
DC/DC Converters	0.00028	0.0056	0.017
Heaters	0.00014	0.0029	0.0086
Horizon Sensors	0.00041	0.0083	0.025
Programmers	0.00089	0.018	0.054
Receivers	0.0015	0.0074	0.018
Sequencers	0.0033	0.016	0.039
Solar Aspect Sensors	0.00055	0.012	0.034
Timers and Clocks	0.00024	0.0048	0.014
Transponders	0.0045	0.033	0.070
Voltage Control Oscillators	0.00075	0.015	0.045

FIGURE 9 - IN-ORBIT FAILURE RATE ESTIMATES AND 90-PERCENT CONFIDENCE INTERVALS FOR SELECTED SPACECRAFT COMPONENTS BASED ON COMBINED DATA SAMPLE

	In-Orbit Failure Rate (Failures/Million Hours)		
	λ_1	$\hat{\lambda}$	λ_2
Batteries	1.3	2.7	5.0
Decoders	0.024	0.48	2.3
Command Distribution Units	0.65	3.7	12.0
Computers	1.8	36.0	166.0
DC/DC Converters	0.62	2.3	5.9
Heaters	0.022	0.43	2.0
Horizon Sensors	4.8	17.0	45.0
Magnetic Tape Units	30.0	40.0	52.0
Motors	0.15	0.79	2.5
Oscillators	0.031	0.54	2.9
Receivers	0.14	0.79	2.5
Regulators, pressure	0.15	4.0	14.0
Regulators, voltage	0.35	1.0	2.4
Telemetry Encoders	4.7	9.5	17.0
Timers and Clocks	3.0	5.6	9.5
Transmitters	1.9	3.4	5.7
Transponders	0.16	3.2	15.0
Vidicon Cameras	4.0	10.0	21.2

FIGURE 10 - IN-ORBIT FAILURE RATE ESTIMATES AND 90-PERCENT CONFIDENCE INTERVALS FOR SELECTED PIECE-PARTS BASED ON COMBINED DATA SAMPLE

	In-Orbit Failure Rate (Failures/Million Hours)		
	λ_1	$\hat{\lambda}$	λ_2
Battery cells	0.0011	0.022	0.10
Capacitors	0.00045	0.00087	0.0015
Diodes	0.00041	0.0012	0.0028
Fuses	0.092	0.33	0.87
Integrated Circuits	0.0038	0.011	0.026
Relays	0.00055	0.011	0.051
Resistors	0.000011	0.00021	0.0020
Solenoids	0.032	0.61	2.9
Switches	0.10	0.37	2.0
Thermistors	0.11	0.28	0.59
Transistors	0.000033	0.00065	0.0031
Traveling Wave Tubes	0.91	1.8	8.5
Tubes, Special Purpose	0.31	6.0	29.0
Geiger Mueller Tubes	5.5	16.0	37.0
Photomultiplier Tubes	0.25	4.6	22.0

FIGURE 11 - FAILURE RATE TRENDS

Piece Part	Failure Rate Reduction Factor
Capacitors	4.5
Diodes	3.3
Resistors	4.2
Transistors	4.2

FIGURE 12 - FAILURE RATE COMPARISON

Piece Part Categories	Failure Rate (in failures per 10^6 hours)			
		Commonly Used Rates		
	Figure 9	Minimum	Maximum	Mil Handbook 217/A
Capacitors	0.00087	0.00079	0.10	0.005
Diodes	0.0012	0.0014	0.20	0.10
Fuses	0.33	0.10	0.50	0.10
Relays	0.011	0.30	1.50	
Resistors	0.00021	0.00024	0.16	0.0033
Solenoids	0.61	0.30	2.5	
Switches, General	0.37	0.023	0.50	
Thermistors	0.28	0.05	0.60	0.30
Transistors	0.00065	0.0004	0.61	0.10
Tubes (Special Purpose)	6.0	0.1	30	

FIGURE 13 - FAILURE RATE STATISTICS FOR SIX SELECTED COMPONENTS

Hardware Element	Failure Rate (Failures/Million Hours)			
	Dormancy	On-Orbit		
	λ_2	λ_1	$\hat{\lambda}$	λ_2
DC/DC Converters	5.7	0.62	2.3	5.9
Magnetic Tape Units	5.5	30.0	40.0	52.0
Transmitters	1.7	1.9	3.4	5.7
Transponders	15.0	0.16	3.2	15.0
Vidicon Cameras	8.4	4.0	10.0	21.2
Switches	1.7	0.10	0.37	2.0

MAINTENANCE STRATEGIES FOR AMBIGUOUS FAULTS

Frank A. Eble
RCA Corporation
Moorestown, New Jersey

Summary

When fault isolation capability is limited to identifying large groups of circuit cards, system restoration times and sparing requirements can be burdensome. The FARO computer program optimizes group replacement strategy, softening the impact of fault ambiguity on operational availability and system support. This paper describes FARO (Fault Ambiguity - Repair Optimization) and shows how to use it effectively.

Introduction

Modern electronic systems are orders of magnitude more complex and densely packaged than their ancestors of just a few years ago. Automated fault isolation to one plug-in circuit card out of thousands may be economically unfeasible, necessitating ambiguous fault indications. Fault ambiguity means that we must be satisfied with the identification of a "fault group" of cards in which a failure has occurred. Since a number of suspect cards usually must be replaced with spares to eliminate a single failure, ambiguous fault sensing incurs system penalties measured in terms of corrective maintenance time and logistic cost.

Circuit complexity rules out manual troubleshooting as a rapid means of isolating the offending card. But maintenance strategies can be devised to cushion the operational and logistic impact of fault ambiguity. Knowing the failure rate of each card, we can employ a probabilistic method to find the culprit in minimum time or with minimum required spares. This paper describes a computerized technique for optimizing fault group replacement strategy and discusses the trade-off considerations associated with its use.

Replacement Strategies

Fault group size and replacement strategy control two important system parameters: Q, the average number of plug-in cards removed from the equipment per failure, and T, the average card-interchange-plus-system-checkout time necessary to restore and verify operation. T influences operational availability as a contributor to system MTTR (Mean-Time-To-Repair); Q and T affect support requirements and life cycle costs. Unfortunately, the twin goals of minimizing Q and T are not always compatible, as we shall see in the following example.

Consider a simple fault group: five cards with equal probabilities of failure. We shall assume that it takes .01 hour to interchange a card with its respective spare, and that the system has a checkout capability of .03 hour. We shall further assume that every suspect card that is removed (a) requires a replacement spare and (b) must be fully tested and certified before being placed in ready spares stock. Two divergent maintenance strategies will be examined:

A. TOTAL FAULT GROUP REPLACEMENT, FOLLOWED BY CHECKOUT:

$$Q = 5 \text{ cards (constant)} \tag{1}$$

$$T = 5(.01) + .03 = .08 \text{ hour (constant)} \tag{2}$$

B. SEQUENTIAL CARD REPLACEMENT, WITH CHECKOUT AFTER EVERY REPLACEMENT; SEQUENCE ENDS WHEN OPERABILITY IS RESTORED:

$$Q = (1+2+3+4+5)/5 = 3 \text{ cards (average)} \tag{3}$$

$$T = 3(.01+.03) = .12 \text{ hour (average)} \tag{4}$$

Strategy A offers a low T value, but incurs maximum Q; Strategy B minimizes Q, at the cost of increased T. The average logistic flows for A and B are compared in Figure 1.

These examples are just two of 16 possible replacement strategies for our hypothetical five-card fault group. Each strategy subdivides the group into a "replacement set" configuration. System checkout is performed after every set replacement; a "go" operability indication terminates the sequence.

Table 1 illustrates all 16 ways of replacing a five-card fault group, following any specified card replacement order. The numbers listed adjacent to each strategy represent sequential replacement set sizes.

Strategy 7, for instance, goes like this: replace the first two cards and perform system checkout; if the fault has not been removed, replace the next card and repeat the checkout operation; if the fault indication is still present, replace the last two cards and perform a final checkout, which should confirm restored system operability.

The uniform-failure-rate example that we have used is an unlikely special case. Real-world differences in card reliabilities allow ordering of replacement sequences on a probabilistic basis to optimize our chances of finding the bad card quickly.

The FARO Program

Our choice of potential replacement strategies for a given card replacement sequence grows exponentially with fault group size. FARO (Fault Ambiguity - Repair Optimization) is a computer program which has been developed to aid the maintenance planner in selecting an optimum strategy. FARO has been written in FORTRAN IV for batch processing on an RCA SPECTRA 70/55 system.

Understanding FARO

The FARO program simulates card removal by sequential replacement sets, which can vary in size from a single card to the entire fault group. Checkout is performed after every set replacement, as described for our five-card example. FARO operates as follows:

Each card of an N-card fault group must be assigned a priority from 1 through N, representing its place in the planned removal sequence. The FARO user will customarily arrange the cards in descending order of failure rate contribution to the group (decreasing probability of having caused the fault signal). Where a card appears in several fault groups, its failure rate contribution to any given group depends on that group's relative probability of detection.

FARO evaluates all possible replacement set configurations for the specified card removal sequence, and calculates an expected Q and T for every configuration, using equations (5) and (6).

$$Q_i = \frac{\sum_{j=1}^{m_i} n_{ij} \sum_{k=j}^{m_i} \sum_{q=1}^{n_{ik}} \lambda_{ikq}}{\sum_{j=1}^{m_i} \sum_{p=1}^{n_{ij}} \lambda_{ijp}} \tag{5}$$

$$T_i = \frac{t_c \sum_{j=1}^{m_i} \sum_{k=j}^{m_i} \sum_{q=1}^{n_{ik}} \lambda_{ikq}}{\sum_{j=1}^{m_i} \sum_{p=1}^{n_{ij}} \lambda_{ijp}} + Q_i\,(t_d + t_i) \tag{6}$$

where
- Q_i = expected number of cards removed for configuration i
- T_i = expected card-interchange-plus-checkout time for configuration i
- m_i = number of replacement sets in configuration i
- n_{ik} = number of cards in replacement set k of configuration i
- n_{ij} = number of cards in replacement set j of configuration i
- λ_{ikq} = failure rate of card q in replacement set k of configuration i
- λ_{ijp} = failure rate of card p in replacement set j of configuration i
- t_c = system checkout time
- t_d = decision time required to match a spare card to an equipment location
- t_i = card interchange time

The λ_{ikq} and λ_{ijp} parameters are expressed in failures per 10^6 hours; all times are expressed in hours.

The effect of fault group and replacement set size on t_d was investigated. Our conclusion: variations in t_d are sufficiently small relative to t_i to be disregarded. Therefore the FARO runs that will be discussed in this paper consider decision time as a constant element of t_i. FARO provides for separate treatment of t_d if warranted by future investigations.

An N-card fault group has 2^{N-1} possible replacement set configurations for any specified card removal sequence. The computer program systematically develops every configuration by assigning values of either 0 or 1 to bit positions located between pairs of adjacent cards in the replacement sequence. A 0 bit value means that both cards are in the same replacement set; a 1 value means that they are in different sets. FARO generates the bit patterns for all binary numbers from 0 through $2^{N-1}-1$, with the least significant bit arbitrarily located between the first and second cards. Every bit pattern defines a unique replacement set configuration.

Computer running time increases sharply as fault group size grows, since 2^{N-1} tests, each involving N-1 bit values, must be performed for a fault group containing N cards. A practical cutoff point for N has been set at 15 cards. Despite this limitation, FARO has been designed to handle fault groups of any size by collecting cards into indivisible "replacement units"; a unit may be composed of any number of cards. Any fault group of 16 or more cards must include at least one multiple-card unit.

Some assumptions of the FARO program:

1. All cards of a fault group are physically located in a common access space. Where portions of the group lie in different drawers, racks, or cabinets, the group can be easily partitioned, with the highest failure-rate portion accessed first. FARO then can be used to optimize replacement strategy within each subgroup.

2. Spares which are inserted in the system and do not correct a fault are left in place. All cards removed from the equipment must undergo the full checkout cycle to prevent the return of suspect cards to the ready spares complement.

3. The program does not deal with failures of backplane wiring or other problems which cannot be eliminated by card substitution. These faults are normally isolated by special maintenance procedures initiated after two successive full fault group replacements have failed to correct the trouble.

Using FARO

FARO accepts fault group size, replacement sequence, card failure rates, interchange time, and checkout time as input data. It calculates as many as five optimization functions $f_n(Q, T)$ for every card replacement strategy applicable to the group, and lists the five best strategies for minimizing each function. Replacement set configuration and computed Q and T values are displayed for every selected strategy. The Q and T associated with total fault group replacement (Strategy A in our original example) are printed out as Q (base) and T (base).

The five optimization functions are listed below:

$f_1(Q,T) = Q$ (Q is the only optimization criterion)
$f_2(Q,T) = Q^2T$ (Q is considered more important than T)
$f_3(Q,T) = QT$ (Q and T are considered equally important)
$f_4(Q,T) = QT^2$ (T is considered more important than Q)
$f_5(Q,T) = T$ (T is the only optimization criterion)

A user of FARO can specify the one optimization function that most closely approximates his estimate of the relative importance of Q and T. Or he can instruct the program to optimize any or all functions for comparison purposes.

Appendix A contains results of sample FARO runs performed on an assumed fault group, with card failure rates as listed in Table 2. Card interchange time is set at .01 hour. Three system checkout times (.02, .06, and .10 hour) and two card/unit arrangements (ten single-card units and five two-card units) are evaluated.

Multiple-card units dramatically reduce program running time: 512 tests for ten one-card units versus only 16 for five two-card units. Yet the sample FARO printouts show that the 512 tests yield only slightly better solutions than the short-cut 16-test approach. Example: the optimum replacement set configuration for function QT in run 1 (checkout time = .02 hour) gives a QT value of 0.58245 for the ten-unit model and 0.59716 for five units.

The FARO "shopping list" of five preferred card replacement strategies per optimization function enables the user to consider the following convenience/efficiency factors in making a selection:

1. Proximity of card locations within sets
2. Standardization of card types within sets
3. Uniformity of set size within group

Figure 2 illustrates a case in point. A certain 7-card fault group is spread over two 20-card nests. Numbers from 1 through 7 designate card replacement priorities, arranged in descending failure rate order.

Let's assume that we have decided to optimize QT^2, and that FARO prints out a preferred strategy of "5,1,1", with "4,3" a close second. For ease of replacement we probably would select "4,3". Or suppose that cards 1, 2, and 3 are of a single type. From a human factors standpoint, perhaps an initial replacement set size of 3 would be preferable.

Summary of FARO Results

Some tentative conclusions can be drawn from the FARO printouts reproduced in Appendix A. Further computer analysis of various fault group sizes and failure rate distributions is necessary to prove generality.

1. Low-Q strategies generally involve more replacement sets than low-T strategies.

2. The lower the system checkout time, the more compatible are the goals of minimum Q and minimum T. With zero checkout time (instantaneous status display when card substitution is performed), the strategy which minimizes Q also minimizes T.

3. As checkout time increases with respect to card interchange time, the disparity between strategies which minimize Q and those which minimize T grows.

4. Even if only T is considered in selecting an optimum card replacement strategy, the resultant Q should be significantly lower than Q (base) - provided that $t_c < 10\, t_i$.

5. Multi-card units yield excellent savings in computer time with only a small loss of optimization: a point to remember when making FARO production runs on hundreds of fault groups in an actual system.

The fault group used in our example has a fairly narrow failure rate spread among its ten cards. Wider failure rate ranges which may be encountered in real fault groups will allow FARO to produce even better Q/Q (base) and T/T (base) payoffs.

FARO uses predicted failure rates to optimize maintenance of newly designed equipment. The resulting replacement strategies should be revised as reliability predictions are updated by field experience. FARO can be incorporated in system monitoring software, with fault group printouts programmed to indicate recommended replacement strategies.

The Q Versus T Problem

Effective utilization of FARO depends on selecting the most appropriate optimization function. This task demands a carefully reasoned judgment as to the relative significance of Q and T. The Q and T parameters influence availability and logistics. Operational availability would appear to depend principally on T - a premise that will be examined in the next section of this paper. Logistic elements (manning, facilities, costs) can be related to T and Q by maintenance analysis. The effect of Q on the logistic support world will be a topic for future study.

The Simulated System

Relationships among Q, T, and system availability are explored below, using an extension of our familiar ten-card fault group. Consider a large shipboard digital processing system. Our simplified system model is composed of 1,000 such groups - 10,000 logic cards in all. Assume that all suspect cards must be returned to a remote location for checkout and repair; that the system is restocked with spares every 30 days; and that a sufficient spares protection level is maintained to support the worst-Q strategy which demands 10 replacement cards per failure.

Spares Sufficiency

The validity of the sufficient-spares assumption must be confirmed. Expected failures per 30-day replenishment cycle are expressed by:

$$\mu_F = \frac{720 \sum_{i=1}^{10} \lambda_i}{10^3} = 27.9 \qquad (7)$$

where λ_i = failures per 10^6 hours for card i

Using the normal approximation to a Poisson probability distribution,

$$\sigma_F \approx \sqrt{27.9} \approx 5.28 \qquad (8)$$

A spares complement of 49 cards per type $(\mu_F + 4\sigma_F)$ should assure greater than 0.9999 probability of surviving a 30-day replenishment cycle even with full fault group replacement for every failure. Since this sparing level represents less than 5 percent of the operational card count, it is judged economically and physically practical to maintain. Therefore the availability analysis of our assumed system will not consider sparing shortages.

Spares protection levels can become critical for low-population, high-failure rate modules which also happen to be (a) costly or (b) bulky to spare. Fortunately, the criticality of such hardware normally justifies unambiguous fault isolation directly to the failed item, reducing Q to the ideal value of one.

With guaranteed spares protection, the Q parameter loses much of its impact on availability. But it does retain some effect via two mechanisms:

1. Cycling of plug-in card connectors

2. Less-than-perfect quality of spares, caused by shelf environment or by damage during transport, testing, or repair

Connector Cycling Effects

The effect on availability of connector remove/replace cycles may be evaluated by applying the following expression[1] for connector failure rate:

$$\lambda_p = \lambda_b \; (\pi_E \cdot \pi_p) + N \; \Sigma_{cyc} \qquad (9)$$

where
λ_p = part application failure rate
λ_b = base failure rate
π_e = environmental factor
π_p = active pin quantity modifier
N = number of active pins
Σ_{cyc} = 0.001 exp (F/100)
F = insertion/withdrawal cycles per 1,000 hours

We shall investigate only the N Σ_{cyc} term; the first term is independent of connector cycling. Two maintenance strategies with widely varying Q values (see Appendix A) will be compared for our simulated system:

1. Full fault group replacement (Q = 10)
2. One-card-at-a-time replacement ($Q \approx 5$)

Card removal/replacement cycles at the equipment location per card of type i per thousand hours are calculated by equation (10) for Strategy 1 and equation (11) for Strategy 2. Results are listed in Table 3.

$$F_{i,1} = 10^{-3} \sum_{j=1}^{10} \lambda_j \qquad (10)$$

$$F_{i,2} = 10^{-3} \sum_{j=i}^{10} \lambda_j \qquad (11)$$

where λ_j = failures per 10^6 hours for card j

Table 3 also tabulates the additional failures per 10^6 hours per card of type i incurred by the additional connector cycling of Strategy 1. Two connector cycles are applied to every card replacement, to allow for card testing at a depot facility. A 70-pin connector is assumed; we make no distinction between card and backplane connector elements for this analysis. The added failure rate for card i is computed by equation (12).

$$\delta\lambda_i = 0.07 \left[\exp\,(2F_{i,1}/100) - \exp\,(2F_{i,2}/100) \right] \qquad (12)$$

Equation (13) calculates the percentage increase in fault group (or system) failure rate produced by employing Strategy 1 in preference to Strategy 2.

$$\frac{100 \sum_{i=1}^{10} \delta\lambda_i}{\sum_{i=1}^{10} \lambda_i} = \frac{0.0273}{38.75} = 0.0007\% \qquad (13)$$

This result shows that the impact of the Q parameter on system availability by way of the N Σ_{cyc} term is negligible.

Spares Quality Effects

Spares quality level influences availability via the occasional faulty spare which requires a second, time-consuming pass through the fault group replacement sequence. We shall develop an expression for operational availability (A_o) which includes the effect of spares defects, based on these assumptions:

1. The incidence of defective spares is directly proportional to Q, and is invariant over all card types for each spares quality level. The latter assumption might be less reasonable for fault groups having wider card failure rate ranges than the example used in our analysis. Further work is necessary to evaluate the extent of spares quality dependence on card failure and/or removal rates.

2. Bad spares are encountered randomly.

3. Completing the fault group replacement sequence without clearing the fault denotes a defective spare. The

sequence is then repeated with a fresh set of spares, replacing single-card sets until the fault is eliminated.

4. Every defective spare encountered is associated with a different failure event. This assumption causes a small negative error in A_o.

5. Once a faulty spare is inserted, all additional spares required as a result of this event are good. This assumption causes a small positive error in A_o.

Availability Calculations

Operational availability of the simulated system is related to Q, T, and spares quality level by equation (14). Spares insufficiency and connector cycling are disregarded, in accordance with our previous conclusions.

$$A_o \approx 1 - 10^{-6} \lambda_s (T + T_1 + QD_sT_2) \qquad (14)$$

where
λ_s = system failures per 10^6 hours
T = average card-replacement-plus-checkout time (from FARO)
T_1 = average time to isolate to the fault group (0.02 hour), obtain spares (0.10 hour), and open and close the equipment cabinet (0.13 hour), for a total time of 0.25 hour
Q = average spares required per failure (from FARO)
D_s = defective spares per total spares (0.01 or less to avoid error)

and T_2, the average additional replacement-plus-checkout time incurred by a bad spare, is approximated by equation (15).

$$T_2 \approx T_{max} - T + T_{log} + \frac{Q}{2}(t_i + t_c) \qquad (15)$$

where
T_{max} = card-replacement-plus-checkout time required to replace entire fault group
T_{log} = logistic time to obtain second set of spares (assumed equal to 0.10 hour)
t_i = card interchange time
t_c = system checkout time

Although equation (14) disregards spares insufficiency it does contain a logistic downtime term (time to acquire spares) which is treated as a constant in this example.

Equation (15) assumes that (a) the expected number of cards, Q, would have been required if the bad spare had not appeared; (b) the remainder of the fault group had to be replaced before the problem could be recognized; and (c) the event occurred halfway through the expected card replacement quantity, Q.

The OARS Simulation

A Monte Carlo computer program, OARS (Operational Availability/Replacement Strategy), has been written as an independent means of calculating A_o. OARS runs made on our simulated 10,000-card system check closely with expected-value calculations performed using equation (14). This correlation confirms the validity of FARO algorithms (5) and (6). Maximum discrepancy between equation (14) and OARS results for the same A_o calculation was approximately 2 parts in 10,000. The expected T_2 values calculated by equation (15) were inputted to OARS; hence the simulation program does not check the accuracy of T_2.

OARS generates an array of real random numbers uniformly distributed on (0,1) for every 24-hour day (see acknowledgement below). It compares these numbers with cumulative Poisson probability values[2] to simulate failure arrivals, tabulates daily downtimes, and computes availabilities over any specified period. OARS also calculates the maximum number of spares of each card type needed during any 30-day replenishment cycle over the simulation period (neglecting spares defects).

Availability as a Function of Q and T

The optimum strategies generated by FARO and compiled in Appendix A were compared using (a) equation (14) and (b) a 10-year OARS simulation. Eighteen trials were made; each trial computed A_o for the optimum FARO strategy corresponding to a specific t_c and f(Q, T). These calculations were performed for three spares quality levels: 1, 0.1, and zero percent defective.

Appendix B shows the results of the OARS simulation. The tabulated results prove that A_o is far more responsive to T than to Q for the cases analyzed - given adequate quantities of spares.

Conclusions

1. FARO offers worthwhile operational and logistic payoffs through optimized replacement strategies for ambiguous fault groups. Production runs covering many groups can be made rapidly, especially if multi-card units are used to conserve computation time.

2. While T influences operational availability via MTTR, the primary impact of Q is on logistic parameters: manpower, facilities, and dollars. The decision to use a particular (Q, T) optimization function will not commonly be made on purely mathematical grounds. The classic trade-off of availability versus cost must still be based largely on judgment factors and external constraints.

3. The OARS simulation program has growth potential for expanded studies involving card testing and repair queues at organizational and depot echelons. We plan to pursue those studies in a continuing search for optimum maintenance strategies.

Acknowledgement

The OARS program generates random numbers by means of a subroutine written by Dr. R. S. Johnson of RCA.

References

1. Proposed MIL-HANDBOOK-217B, "Reliability Prediction," April 1970.
2. "Tables of the Individual and Cumulative Terms of Poisson Distribution," General Electric Company Defense Systems Department, 1962.

TABLE 1. REPLACEMENT SET CONFIGURATIONS FOR FIVE-CARD FAULT GROUP

Configuration No.	Set 1	Set 2	Set 3	Set 4	Set 5
1	5	-	-	-	-
2	1	4	-	-	-
3	2	3	-	-	-
4	1	1	3	-	-
5	3	2	-	-	-
6	1	2	2	-	-
7	2	1	2	-	-
8	1	1	1	2	-
9	4	1	-	-	-
10	1	3	1	-	-
11	2	2	1	-	-
12	1	1	2	1	-
13	3	1	1	-	-
14	1	2	1	1	-
15	2	1	1	1	-
16	1	1	1	1	1

TABLE 2. CARD FAILURE RATES FOR ASSUMED FAULT GROUP

Card Type	Card Failures per 10^6 Hours
1	5.00
2	4.75
3	4.50
4	4.25
5	4.00
6	3.75
7	3.50
8	3.25
9	3.00
10	2.75

TABLE 3. EFFECT OF CONNECTOR CYCLING ON CARD FAILURE RATE

Card Type i	F_i (Cycles/1,000 Hrs.) Strategy 1	Strategy 2	$\delta\lambda_i$ (Failures/ 10^6 Hrs.)
1	0.03875	0.03875	0.
2	0.03875	0.03375	0.000007
3	0.03875	0.02900	0.000013
4	0.03875	0.02450	0.000020
5	0.03875	0.02025	0.000026
6	0.03875	0.01625	0.000032
7	0.03875	0.01250	0.000037
8	0.03875	0.00900	0.000041
9	0.03875	0.00575	0.000046
10	0.03875	0.00275	0.000051
			$\Sigma\delta\lambda_i$ = 0.000273

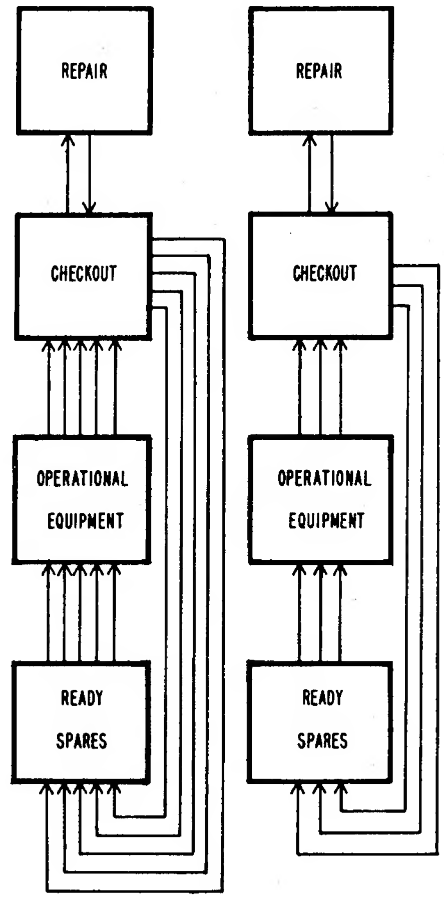


FIGURE 1. AVERAGE LOGISTIC FLOWS FOR FIVE-CARD FAULT GROUP

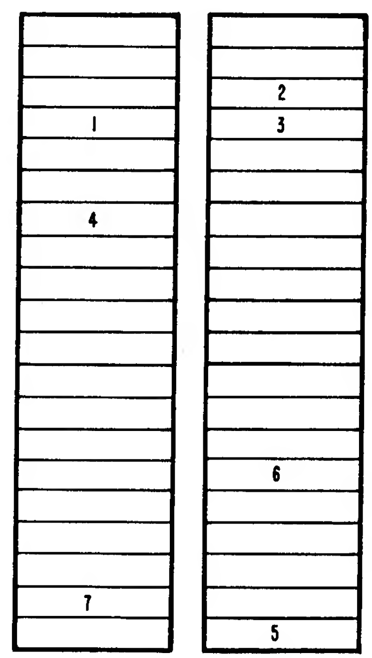


FIGURE 2. GEOGRAPHY OF A SEVEN-CARD FAULT GROUP

APPENDIX A - SAMPLE FARO PRINTOUTS

GROUP	RUN	INTERCH.	CHECKOUT	DECISION	BASE Q	BASE T
1	1	0,0100	0,0200	0,0000	10,0000	0,1200

RANK	Q	Q	T	TEST	SETS	CARDS PER SET
1	4,96774	4,9677	0,1490	512	10	1 1 1 1 1 1 1 1 1 1
2	5,04516	5,0452	0,1484	256	9	1 1 1 1 1 1 1 1 2
3	5,05161	5,0516	0,1469	384	9	1 1 1 1 1 1 1 2 1
4	5,05806	5,0581	0,1453	448	9	1 1 1 1 1 1 2 1 1
5	5,06452	5,0645	0,1435	480	9	1 1 1 1 1 2 1 1 1

RANK	QQT	Q	T	TEST	SETS	CARDS PER SET
1	3,22360	5,3161	0,1141	491	7	2 2 2 1 1 1 1
2	3,22469	5,4065	0,1103	427	6	2 2 2 2 1 1
3	3,25881	5,5548	0,1056	469	6	3 2 2 1 1 1
4	3,26403	5,4000	0,1119	363	6	2 2 2 1 2 1
5	3,26853	5,3097	0,1159	475	7	2 2 1 2 1 1 1

RANK	QT	Q	T	TEST	SETS	CARDS PER SET
1	0,58245	5,7613	0,1011	421	5	3 3 2 1 1
2	0,58351	5,6387	0,1035	341	5	3 2 2 2 1
3	0,58651	5,8387	0,1005	165	4	3 3 2 2
4	0,58666	5,5548	0,1056	469	6	3 2 2 1 1 1
5	0,59049	5,7419	0,1028	405	5	3 2 3 1 1

RANK	QTT	Q	T	TEST	SETS	CARDS PER SET
1	0,05808	6,1032	0,0975	329	4	4 3 2 1
2	0,05888	5,7613	0,1011	421	5	3 3 2 1 1
3	0,05892	5,8387	0,1005	165	4	3 3 2 2
4	0,05920	5,9355	0,0999	293	4	3 3 3 1
5	0,05956	5,9097	0,1004	425	5	4 2 2 1 1

RANK	T	Q	T	TEST	SETS	CARDS PER SET
1	0,09690	6,5548	0,0969	145	3	5 3 2
2	0,09729	6,7484	0,0973	273	3	5 4 1
3	0,09729	6,3871	0,0973	137	3	4 4 2
4	0,09755	6,1032	0,0975	329	4	4 3 2 1
5	0,09755	6,4774	0,0975	401	4	5 3 1 1

GROUP	RUN	INTERCH.	CHECKOUT	DECISION	BASE Q	BASE T
1	2	0,0100	0,0600	0,0000	10,0000	0,1600

RANK	Q	Q	T	TEST	SETS	CARDS PER SET
1	4,96774	4,9677	0,3477	512	10	1 1 1 1 1 1 1 1 1 1
2	5,04516	5,0452	0,3443	256	9	1 1 1 1 1 1 1 1 2
3	5,05161	5,0516	0,3397	384	9	1 1 1 1 1 1 1 2 1
4	5,05806	5,0581	0,3347	448	9	1 1 1 1 1 1 2 1 1
5	5,06452	5,0645	0,3294	480	9	1 1 1 1 1 2 1 1 1

RANK	QQT	Q	T	TEST	SETS	CARDS PER SET
1	6,24232	5,7613	0,1881	421	5	3 3 2 1 1
2	6,28516	5,6387	0,1977	341	5	3 2 2 2 1
3	6,29244	5,8387	0,1846	165	4	3 3 2 2
4	6,34841	5,5548	0,2057	469	6	3 2 2 1 1 1
5	6,35402	6,1032	0,1706	329	4	4 3 2 1

RANK	QT	Q	T	TEST	SETS	CARDS PER SET
1	1,04109	6,1032	0,1706	329	4	4 3 2 1
2	1,04624	6,5548	0,1596	145	3	5 3 2
3	1,04831	6,3871	0,1641	137	3	4 4 2
4	1,05203	6,2645	0,1679	73	3	4 3 3
5	1,05645	6,4774	0,1631	401	4	5 3 1 1

RANK	QTT	Q	T	TEST	SETS	CARDS PER SET
1	0,16614	6,7484	0,1569	273	3	5 4 1
2	0,16693	7,0387	0,1540	289	3	6 3 1
3	0,16699	6,5548	0,1596	145	3	5 3 2
4	0,16901	7,2903	0,1523	33	2	6 4
5	0,17206	6,3871	0,1641	137	3	4 4 2

RANK	T	Q	T	TEST	SETS	CARDS PER SET
1	0,15090	7,6968	0,1509	65	2	7 3
2	0,15187	8,2968	0,1519	129	2	8 2
3	0,15226	7,2903	0,1523	33	2	6 4
4	0,15355	7,5355	0,1535	321	3	7 2 1
5	0,15400	7,0387	0,1540	289	3	6 3 1

GROUP	RUN	INTERCH.	CHECKOUT	DECISION	BASE Q	BASE T
1	3	0,0100	0,1000	0,0000	10,0000	0,2000

RANK	Q	Q	T	TEST	SETS	CARDS PER SET
1	4,96774	4,9677	0,5465	512	10	1 1 1 1 1 1 1 1 1 1
2	5,04516	5,0452	0,5401	256	9	1 1 1 1 1 1 1 1 2
3	5,05161	5,0516	0,5325	384	9	1 1 1 1 1 1 1 2 1
4	5,05806	5,0581	0,5241	448	9	1 1 1 1 1 1 2 1 1
5	5,06452	5,0645	0,5152	480	9	1 1 1 1 1 2 1 1 1

RANK	QQT	Q	T	TEST	SETS	CARDS PER SET
1	9,07442	6,1032	0,2436	329	4	4 3 2 1
2	9,12899	5,7613	0,2750	421	5	3 3 2 1 1
3	9,16045	5,8387	0,2687	165	4	3 3 2 2
4	9,22797	5,9355	0,2619	293	4	3 3 3 1
5	9,27408	5,9097	0,2655	425	5	4 2 2 1 1

RANK	QT	Q	T	TEST	SETS	CARDS PER SET
1	1,45729	6,5548	0,2223	145	3	5 3 2
2	1,46113	6,7484	0,2165	273	3	5 4 1
3	1,47521	6,3871	0,2310	137	3	4 4 2
4	1,47631	7,0387	0,2097	289	3	6 3 1
5	1,48103	6,4774	0,2286	401	4	5 3 1 1

RANK	QTT	Q	T	TEST	SETS	CARDS PER SET
1	0,30686	7,2903	0,2052	33	2	6 4
2	0,30847	7,6968	0,2002	65	2	7 3
3	0,30964	7,0387	0,2097	289	3	6 3 1
4	0,31636	6,7484	0,2165	273	3	5 4 1
5	0,31877	7,5355	0,2057	321	3	7 2 1

RANK	T	Q	T	TEST	SETS	CARDS PER SET
1	0,19781	8,2968	0,1978	129	2	8 2
2	0,19781	9,0710	0,1978	257	2	9 1
3	0,20000	10,0000	0,2000	1	1	10
4	0,20019	7,6968	0,2002	65	2	7 3
5	0,20413	8,2194	0,2041	385	3	8 1 1

GROUP	RUN	INTERCH.	CHECKOUT	DECISION	BASE Q	BASE T
2	1	0,0100	0,0200	0,0000	10,0000	0,1200

RANK	Q	Q	T	TEST	SETS	CARDS PER SET
1	5,48387	5,4839	0,1097	16	5	2 2 2 2 2
2	5,83226	5,8323	0,1102	8	4	2 2 2 4
3	5,88387	5,8839	0,1072	12	4	2 2 4 2
4	5,93548	5,9355	0,1037	14	4	2 4 2 2
5	5,98710	5,9871	0,0997	15	4	4 2 2 2

RANK	QQT	Q	T	TEST	SETS	CARDS PER SET
1	3,29831	5,4839	0,1097	16	5	2 2 2 2 2
2	3,57528	5,9871	0,0997	15	4	4 2 2 2
3	3,65482	5,9355	0,1037	14	4	2 4 2 2
4	3,71215	5,8839	0,1072	12	4	2 2 4 2
5	3,74825	5,8323	0,1102	8	4	2 2 2 4

RANK	QT	Q	T	TEST	SETS	CARDS PER SET
1	0,59716	5,9871	0,0997	15	4	4 2 2 2
2	0,60146	5,4839	0,1097	16	5	2 2 2 2 2
3	0,61576	5,9355	0,1037	14	4	2 4 2 2
4	0,62140	6,3871	0,0973	11	3	4 4 2
5	0,63090	5,8839	0,1072	12	4	2 2 4 2

RANK	QTT	Q	T	TEST	SETS	CARDS PER SET
1	0,05956	5,9871	0,0997	15	4	4 2 2 2
2	0,06046	6,3871	0,0973	11	3	4 4 2
3	0,06368	6,3355	0,1003	7	3	4 2 4
4	0,06388	5,9355	0,1037	14	4	2 4 2 2
5	0,06597	5,4839	0,1097	16	5	2 2 2 2 2

RANK	T	Q	T	TEST	SETS	CARDS PER SET
1	0,09729	6,3871	0,0973	11	3	4 4 2
2	0,09884	6,9419	0,0988	13	3	6 2 2
3	0,09935	7,2903	0,0994	5	2	6 4
4	0,09974	5,9871	0,0997	15	4	4 2 2 2
5	0,10026	6,3355	0,1003	7	3	4 2 4

GROUP	RUN	INTERCH.	CHECKOUT	DECISION	BASE Q	BASE T
2	2	0,0100	0,0600	0,0000	10,0000	0,1600

RANK	Q	Q	T	TEST	SETS	CARDS PER SET
1	5,48387	5,4839	0,2194	16	5	2 2 2 2 2
2	5,83226	5,8323	0,2139	8	4	2 2 2 4
3	5,88387	5,8839	0,2040	12	4	2 2 4 2
4	5,93548	5,9355	0,1925	14	4	2 4 2 2
5	5,98710	5,9871	0,1795	15	4	4 2 2 2

RANK	QQT	Q	T	TEST	SETS	CARDS PER SET
1	6,43365	5,9871	0,1795	15	4	4 2 2 2
2	6,59662	5,4839	0,2194	16	5	2 2 2 2 2
3	6,69564	6,3871	0,1641	11	3	4 4 2
4	6,78233	5,9355	0,1925	14	4	2 4 2 2
5	6,98666	6,3355	0,1741	7	3	4 2 4

RANK	QT	Q	T	TEST	SETS	CARDS PER SET
1	1,04831	6,3871	0,1641	11	3	4 4 2
2	1,07459	5,9871	0,1795	15	4	4 2 2 2
3	1,09459	6,9419	0,1577	13	3	6 2 2
4	1,10278	6,3355	0,1741	7	3	4 2 4
5	1,11001	7,2903	0,1523	5	2	6 4

RANK	QTT	Q	T	TEST	SETS	CARDS PER SET
1	0,16901	7,2903	0,1523	5	2	6 4
2	0,17206	6,3871	0,1641	11	3	4 4 2
3	0,17259	6,9419	0,1577	13	3	6 2 2
4	0,18891	7,1355	0,1627	3	2	4 6
5	0,19136	8,2968	0,1519	9	2	8 2

RANK	T	Q	T	TEST	SETS	CARDS PER SET
1	0,15187	8,2968	0,1519	9	2	8 2
2	0,15226	7,2903	0,1523	5	2	6 4
3	0,15768	6,9419	0,1577	13	3	6 2 2
4	0,16000	10,0000	0,1600	1	1	10
5	0,16271	7,1355	0,1627	3	2	4 6

GROUP	RUN	INTERCH.	CHECKOUT	DECISION	BASE Q	BASE T
2	3	0,0100	0,1000	0,0000	10,0000	0,2000

RANK	Q	Q	T	TEST	SETS	CARDS PER SET
1	5,48387	5,4839	0,3290	16	5	2 2 2 2 2
2	5,83226	5,8323	0,3177	8	4	2 2 2 4
3	5,88387	5,8839	0,3008	12	4	2 2 4 2
4	5,93548	5,9355	0,2813	14	4	2 4 2 2
5	5,98710	5,9871	0,2592	15	4	4 2 2 2

RANK	QQT	Q	T	TEST	SETS	CARDS PER SET
1	9,29203	5,9871	0,2592	15	4	4 2 2 2
2	9,42233	6,3871	0,2310	11	3	4 4 2
3	9,89493	5,4839	0,3290	16	5	2 2 2 2 2
4	9,90984	5,9355	0,2813	14	4	2 4 2 2
5	9,94913	6,3355	0,2479	7	3	4 2 4

RANK	QT	Q	T	TEST	SETS	CARDS PER SET
1	1,47521	6,3871	0,2310	11	3	4 4 2
2	1,49569	7,2903	0,2052	5	2	6 4
3	1,50304	6,9419	0,2165	13	3	6 2 2
4	1,55201	5,9871	0,2592	15	4	4 2 2 2
5	1,57038	6,3355	0,2479	7	3	4 2 4

RANK	QTT	Q	T	TEST	SETS	CARDS PER SET
1	0,30686	7,2903	0,2052	5	2	6 4
2	0,32463	8,2968	0,1978	9	2	8 2
3	0,32543	6,9419	0,2165	13	3	6 2 2
4	0,34073	6,3871	0,2310	11	3	4 4 2
5	0,35679	7,1355	0,2236	3	2	4 6

RANK	T	Q	T	TEST	SETS	CARDS PER SET
1	0,19781	8,2968	0,1978	9	2	8 2
2	0,20000	10,0000	0,2000	1	1	10
3	0,20516	7,2903	0,2052	5	2	6 4
4	0,21652	6,9419	0,2165	13	3	6 2 2
5	0,22361	7,1355	0,2236	3	2	4 6

APPENDIX B - RESULTS OF A TEN-YEAR OARS SIMULATION

TRIAL	t_i	t_c	OPT. FNCT.	SETS	CARDS PER SET	Q	T	AVAILABILITY FOR (%) DEFECTIVE SPARES (1.0)	(0.1)	(0)	CARDS REQUIRED FOR 0% DEF. SPARES 10-YR TOTAL	30-DAY MAXIMUM BY CARD TYPE 1	2	3	4	5	6	7	8	9	10
1	.01	.02	Q	10	1 1 1 1 1 1 1 1 1 1	4.9677	.1490	.983822	.984414	.984463	16770	47	44	40	32	27	23	19	16	10	5
2	.01	.02	Q^2T	7	2 2 2 1 1 1 1	5.3161	.1141	.985182	.985788	.985835	17930	47	47	40	40	27	27	19	16	10	5
3	.01	.02	QT	5	3 3 2 1 1	5.7613	.1101	.985703	.986300	.986347	19402	47	47	47	32	32	32	19	19	10	5
4	.01	.02	QT^2	4	4 3 2 1	6.1032	.0975	.985834*	.986446	.986486	20549	47	47	47	47	27	27	27	16	16	5
5	.01	.02	T	3	5 3 2	6.5548	.0969	.985830	.986459*	.986507*	22080	47	47	47	47	47	23	23	23	10	10
6	.01	.02	BASE	1	10	10.0000	.1200	.984623	.985546	.985622	33580	47	47	47	47	47	47	47	47	47	47
7	.01	.06	Q	10	1 1 1 1 1 1 1 1 1 1	4.9677	.3477	.975475	.976605	.976699	16770	47	44	40	32	27	23	19	16	10	5
8	.01	.06	Q^2T	5	3 3 2 1 1	5.7613	.1881	.981828	.982879	.982961	19402	47	47	47	32	32	32	19	19	10	5
9	.01	.06	QT	4	4 3 2 1	6.1032	.1706	.982482	.983571	.983643	20549	47	47	47	47	27	27	27	16	16	5
10	.01	.06	QT^2	3	5 4 1	6.7484	.1569	.982924	.984085	.984170	22744	47	47	47	47	47	23	23	23	23	5
11	.01	.06	T	2	7 3	7.6968	.1509	.983133*	.984340*	.984421*	25846	47	47	47	47	47	47	47	16	16	16
12	.01	.06	BASE	1	10	10.0000	.1600	.982271	.983932	.984067	33580	47	47	47	47	47	47	47	47	47	47
13	.01	.10	Q	10	1 1 1 1 1 1 1 1 1 1	4.9677	.5465	.967127	.968796	.968936	16770	47	44	40	32	27	23	19	16	10	5
14	.01	.10	Q^2T	4	4 3 2 1	6.1032	.2436	.979132	.980697	.980801	20549	47	47	47	47	27	27	27	16	16	5
15	.01	.10	QT	3	5 3 2	6.5548	.2223	.979945	.981493	.981612	22080	47	47	47	47	47	23	23	23	10	10
16	.01	.10	QT^2	2	6 4	7.2903	.2052	.980596*	.982165	.982290	24536	47	47	47	47	47	47	19	19	19	19
17	.01	.10	T	2	8 2	8.2968	.1978	.980438	.982455*	.982600*	27858	47	47	47	47	47	47	47	47	10	10
18	.01	.10	BASE	1	10	10.0000	.2000	.979917	.982316	.982512	33580	47	47	47	47	47	47	47	47	47	47

*optimum strategy

MAINTENANCE FLOAT FOR SMALL FLEET SIZES

by

Todd E. Stevenson
US Army Test and Evaluation Command
Aberdeen Proving Ground, Maryland

and

Dr. Roger J. McNichols
Texas A&M University
Red River Army Depot
Texarkana, Texas

Summary

The size of a maintenance float needs to be optimized to provide the greatest fleet capability for the least expenditure. This paper presents a method for choosing a maintenance float size based on the anticipation repair rate and equipment failure rate, with allowances for varying these rates to study a wide variety of possible situations.

Introduction

When a unit of a fleet ceases to perform its function due to failure, it becomes necessary to have a unit on hand as a replacement. If there is no spare unit available, the fleet is forced to operate with less than its original number until the failed unit is returned to service. A lack of "full force capability" can be especially detrimental to a small fleet. For example, six aircraft may have to do the work of seven when one of the fleet of seven is down for repairs. A requirement for a number of spare machines to assure a full force is clearly indicated. This article will refer to this pool of spares as the maintenance float.

In a real-world situation, there are usually two major factions competing to determine the size of the float. Comprising the first faction are the fleet managers and operators, who want the float as large as possible to insure a high level of full force capability. Those who must approve the cost of the system, and therefore want as few float items as possible, comprise the second faction. The float size is fixed when the two factions compromise on the number of float items that will give a good full force capability for an agreeable price.

Model

The maintenance float problem can be looked at as essentially a closed loop queue. Queuing theory will provide an algorithm whereby the probabilities for the number of units in the float and in the service facility at any one time can be calculated from the distributions of the times to failure and the times to repair. With these probabilities, the number of units needed to assure a level of full force capability can be determined.

As shown in Figure 1, the fleet is depicted as being of some finite size N. This fleet size is constrained to a definite number N so that the solution procedure can reflect the outcome of operating at less than full force capability; that is, operation with any number less than N units in the fleet.

The number of units in the maintenance float is represented by the letter K. Note that a maintenance float unit can be in any one of four positions: in the float itself waiting to substitute for a failed unit, in the fleet as an operating unit, in the queue waiting for service, or under repair in the service facility. Thus the maintenance float units cycle throughout the system; when there is a large number of units in the service facility and the queue, the maintenance float itself may have relatively few units.

The fleet is operating at its full force size only as long as there are K or less units in the combined queue and service facility. Note that the fleet size can only decrease from its full force size N when two conditions are present simultaneously; first, there must be K units in the combined service facility and queue, and second, there must be another failure before any repair activities are completed in the service facility. Thus K+1 units are in the combined service facility and queue, and N-1 left operating in the field.

The unit failure rate is represented by $\lambda(i)$, where the argument i represents the number of units in the combined service facility and queue. At any point in time the fleet failure rate will be the number of units operating in the fleet times the unit failure rate.

It may be appropriate in certain situations to assume that the unit service rate changes as a function of the number of units in the repair facility. For example, it may be the practice to allow overtime in the maintenance facility when the queue of units waiting for service grows beyond a predetermined point. To accommodate this assumption, the repair rate is assumed to be a function, $\mu(i)$, of the total number of units either waiting for repair or under repair.

The solution will consist mainly of solving the steady state equations for the probabilities that any given number of units will be in the repair facility at any one time. These probabilities form a set denoted by P(i), the probability that there are i units in the service facility and queue. These probabilities will result from the solution of the steady state equations describing the operation of the queue and service facility. For the "steady state", the probability of decreasing the number in the queue and service facility is equal to the probability of increasing that number. With this fact, the following development is possible.

The net change over time in the probability that there are zero units in the queue and service facility can be expressed as:

$$\frac{dP(0)}{dt} = \mu(1)P(1) - N(0)\lambda(0)P(0) = 0. \tag{1}$$

In this equation, the expression $\mu(1)P(1)$ stands for the net shift into the state of a population of zero: that is, the probability that the population is one, times the repair rate when the population is at one, $\mu(1)$. The quantity $N(0)\lambda(0)P(0)$ expresses the net shift out of a population of zero: that is, the probability that the population is zero times the arrival rate, $N(0)\lambda(0)$. This arrival rate is the per unit failure rate $\lambda(0)$ times the number of units operating in the field $N(0)$.

The equation for P(1) includes the respective shifts into and out of the population of one. There are now four terms on the right hand side, representing the fact that there can be a repair or a failure when the popula-

tion is one, or there can be a failure at a population zero, or a repair at a population of two. Thus,

$$\frac{dP(1)}{dt} = -\mu(1)P(1) + N(0)\lambda(0)P(0) - N(1)\lambda(1)P(1) + \mu(2)P(2) = 0. \qquad (2)$$

Note that the first two terms can be dropped out, since they are known to add to zero from Equation 1.

Rearranging the terms leads to a recurring relationship for any two probabilities for the population of the service facility and queue;

$$P(i) = \frac{N(i-1)\lambda(i-1)}{\mu(i)} P(i-1). \qquad (3)$$

Using this relationship, every probability can be expressed as a constant times the probability, P(0), of zero units in the service facility and queue.

Since all the probabilities must add to one, the summation reduces to P(0) times a constant representing the sum of all these aforementioned ratios, or:

$$\sum_{i=0}^{N+K} P(i) = 1.0 = P(0)\ [1 + \frac{N(0)\lambda(0)}{\mu(1)} + \frac{N(0)\lambda(0)N(1)\lambda(1)}{\mu(1)\mu(2)} + \ldots]. \qquad (4)$$

A division rearranges Equation 4 and gives a numerical value for P(0):

$$P(0) = \frac{1.0}{1 + \sum_{j=1}^{N+K} \prod_{i=1}^{j} \frac{N(i-1)\lambda(i-1)}{\mu(i)}} \qquad (5)$$

Using Equation 3, together with the value of P(0) from Equation 5, the numerical values of each probability can be calculated.

The percent of time that the fleet will be at full force can now be determined by adding the probabilities that the population of the service facility and queue will be less than or equal to the maintenance float size, K. This full force capability can be expressed as:

$$FFC = \sum_{i=0}^{K} P(i). \qquad (6)$$

By varying the float size K, the optimum number of float units can be found for a given level of full force capability. Conversely, the capability level for a given system can be determined.

Example:

As an interesting (but non-typical) example, let the service rate decrease as a function of the combined number of units in the service facility and queue. The arrival for service rate is dependent on the number of units actually operating in the field. Starting with the assumption of two maintenance float units, the solution proceeds as shown in the upper half of Table I. From the information in Table I, the probability that the fleet will be at full force can be computed by:

$$\text{Full Force Capability} = P(0) + P(1) + P(2) = .84947$$

Suppose that the capability of .84947 does not meet the required specification for the fleet, and that it has been decided to increase the float size to four. From the lower half of Table I, The Full Force Capability becomes the sum of five probabilities:

$$FFC = P(0) + P(1) + P(2) + P(3) + P(4) = .71355$$

Interestingly for this case, when additional units are added to the float, the Full Force Capability has actually dropped. Intuitively, this seems to be highly unreasonable. It seems at first that the more units on hand for use, the higher the Full Force Capability should be. However, this situation can be analyzed by looking at the tables of the expected number of units in the field versus the expected number of units in the service facility and queue. Table II shows the expected number of units in the field.

The calculations of Table II imply that for this system, as the float size increases, the expected value of the number of units in the field decreases. With the increase from two to four float units, the expected value of the field population drops from 3.668 to 3.368. Essentially, the extra float units in this example increase the size of the queue rather than adding to the expected value of the field size, thus causing the Full Force Capability to drop. In particular, the probability of being in a given state does not continuously decrease as the state number increases. When the float has four units, the probabilities shift enough that the advantage of having more float units is overridden by the probability that more units will be in the queue.

Calculations made with the same set of failure and repair rates produce the results shown in Figure 2, which shows the Full Force Capability as a function of the number of units in the maintenance float.

Probably the most important application of the float size information is in the fleet design stage when the designer needs to know the number of extra units required to maintain a specified Full Force Capability. An associated use for the float size information comes from the fact that the analyst now has a means for determining the cost of increasing the Full Force Capability for a fleet already in the field. Again using the methods of this paper, the fleet designer can evaluate the costs or savings resultant from a change in the service or failure rates.

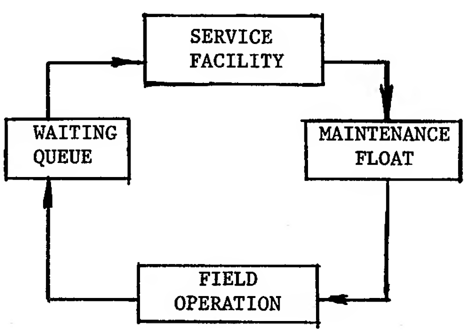
The heart of the analysis lies in the solving of a set of simultaneous equations for the probabilities that certain numbers of units will be in the service facility and queue. An analog computer can be programmed to solve the equations, and to thus provide an insight into the time varying behavior of the system. Figure 3 shows a typical analog diagram for a fleet with three field units and two float units. The output of this program is shown in Figure 4, which illustrates the time dependent behavior of the various probabilities as well as the time dependent form of the Full Force Capability. Repeated use of the analog technique has confirmed that the fleet-float system does indeed come to rest at the steady state values predicted by the analytical solution of the steady state equations.

Conclusion

In conclusion, it can be said that this method provides a mathematically sound answer to the maintenance float question. Further, the method is versatile in that the designer needs only to have an estimate of the repair and failure rates to arrive at a solution. Considering the cost of float units such as aircraft or motor vehicles, this method has a potential of saving the user large sums of money, while only costing him the amount of time necessary to evaluate a few sets of fundamental equations.

References

1. Churchman, C. West, Russel L. Ackoff, and E. Leonard Arnoff Introduction to Operations Research. New York, N.Y.: John Wiley and Sons, Inc., 1957.
2. Cruon, R., A. Rougerie, and C. Van de Casteele "Etude D'un Circuit de Materiels Reparables," Revue Francaise d'Informatique et de Recherch Operationelle, Vol. 3, Nov. 1969.
3. Department of Defense Military Standard, Maintainability Program Requirements, MIL-Std-470, 21 March 1966.
4. Faulkner, J.A. "The Use of Closed Queues in the Deployment of Coal-Face Machinery, "Operations Research Quarterly, Vol. 19, No. 1, March 1968.
5. Hillier, Frederick S. and Gerald J. Lieberman Introduction to Operations Research. San Francisco, Cal.: Holden-Day Inc., 1967.
6. Levine, Boris "Estimating Maintenance Float Factors on the Basis of Reliability Theory," Industrial Quality Control, Feb. 1965.
7. Morris, William T. Analysis for Materials Handling Management. Homewood, Ill.: Richard D. Irwin, Inc., 1962.
8. Saaty, Thomas L. Elements of Queueing Theory. New York, N.Y.: McGraw-Hill Book Company, Inc., 1961.
9. Saaty, Thomas L. Mathematical Methods of Operations Research. New York, N.Y.: McGraw-Hill Book Company, Inc., 1959.
10. Sahu, K.C. and K.C. Sharma "Determination of Optimal Rotating Float in a Closed Loop System: A Case Study," The International Journal of Production Research, Vol. 8, No. 3, 1970.
11. Sasieni, Maurice, Arthur Yaspan, and Lawrence Friedman Operations Research Methods and Problems. New York, N.Y.: John Wiley and Sons, Inc., 1959.
12. Taylor, J. and R.R.P. Jackson "An Application of the Birth and Death Process to the Provision of Spare Machines," Operational Research Quarterly, Vol 5, 1954.
13. Todaro, John B. "Optimization of System Design by Spare or Repair Decision Through Dynamic Programming," Unpublished Research Project, Texas A&M University.
14. U.S. Army Regulation, Maintenance Float Support of Army Materiel, AR 750-19, 24 April 1970.
15. White, Robert Dale "An Optimization Approach to the Determination of Maintenance Float Requirements," Unpublished Research Project, Texas A&M University, December, 1967.



INITIAL STATES:
SERVICE FACILITY....0
WAITING QUEUE.......0
FIELD OPERATIONS....N
MAINTENANCE FLOAT...K

FIGURE 1 THE FLEET MODEL

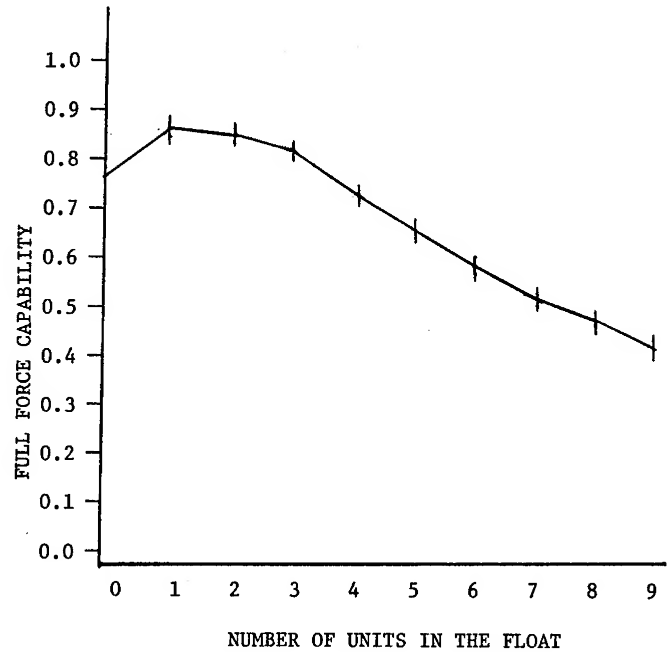


FIGURE 2 FULL FORCE CAPABILITY VERSUS FLOAT SIZE

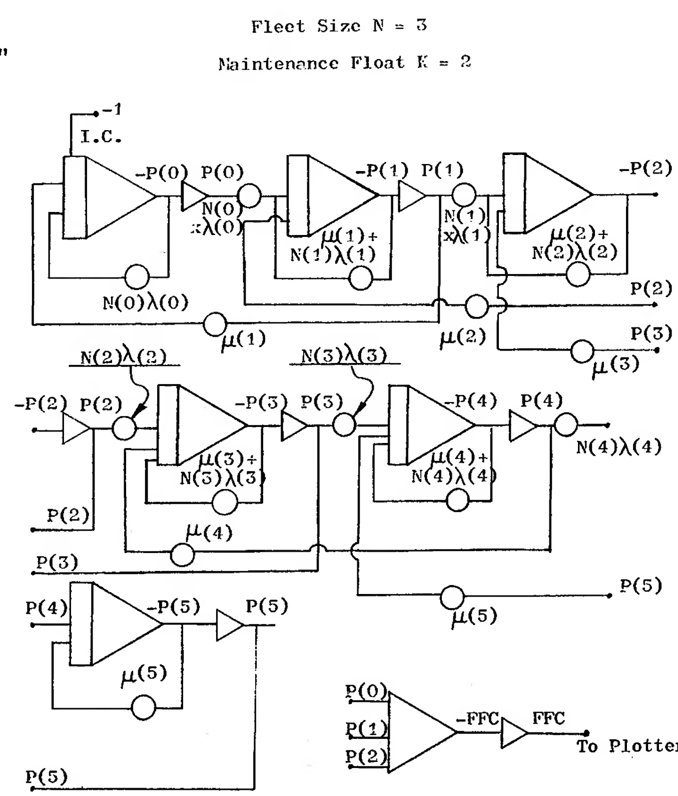


FIGURE 3 EXAMPLE OF AN ANALOG WIRING DIAGRAM

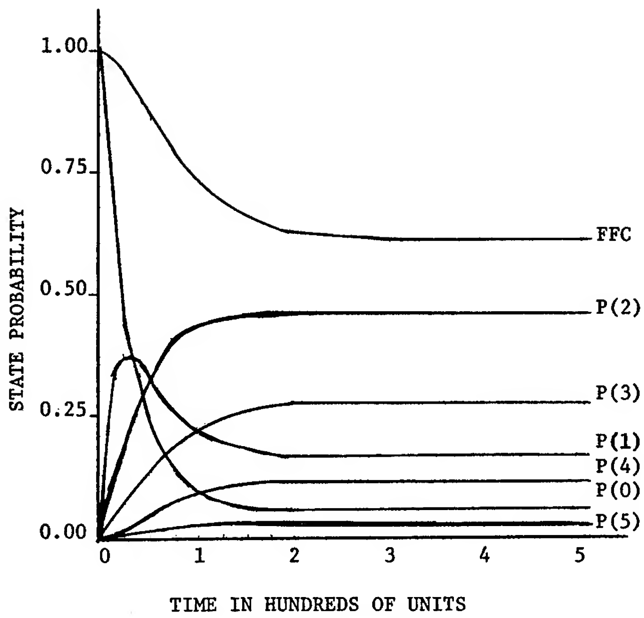


FIGURE 4 EXAMPLE ANALOG OUTPUT

TABLE I

FIRST CASE FOR EXAMPLE FLEET
FLEET SIZE N = 4
MAINTENANCE FLOAT SIZE K = 2

P(0) =		= 1.0	P(0) = .66365
P(1) =	(.40/2.0) P(0)	= .20	P(0) = .13273
P(2) =	(.40/1.0) P(1)	= .08	P(0) = .05309
P(3) =	(.40/0.5) P(2)	= .064	P(0) = .04247
P(4) =	(.30/.25) P(3)	= .0768	P(0) = .05097
P(5) =	(.20/.25) P(4)	= .06144	P(0) = .04078
P(6) =	(.10/.25) P(5)	= .024576	P(0) = .01631
1.0 =		= 1.506816	P(0) =1.00000

FFC = P(0) + P(1) + P(2) = .84947

SECOND CASE FOR EXAMPLE FLEET
FLEET SIZE N = 4
MAINTENANCE FLOAT SIZE K = 4

P(0) =		= 1.0	P(0) = .49333
P(1) =	(.40/2.0) P(0)	= .20	P(0) = .09867
P(2) =	(.40/1.0) P(1)	= .08	P(0) = .03947
P(3) =	(.40/0.5) P(2)	= .064	P(0) = .03157
P(4) =	(.40/.25) P(3)	= .1024	P(0) = .05052
P(5) =	(.40/.25) P(4)	= .16384	P(0) = .08083
P(6) =	(.30/.25) P(5)	= .196608	P(0) = .09699
P(7) =	(.20/.25) P(6)	= .1572864	P(0) = .07759
P(8) =	(.10/.25) P(7)	= .06291456	P(0) = .03103
1.0 =		= 2.02704896	P(0) =1.00000

FFC = P(0) + P(1) + P(2) + P(3) + P(4) = .71355

TABLE II EXPECTED VALUES FOR TWO MAINTENANCE FLOAT SIZES

FIRST CASE
FLEET SIZE N = 4
MAINTENANCE FLOAT SIZE K = 2

i	P(i)	NO. OF UNITS IN FIELD	EXPECTED NO. OF UNITS IN SERVICE AND QUEUE	EXPECTED NO. OF UNITS IN FIELD
0	.664	4	0.000	2.655*
1	.133	4	.133	.531*
2	.053	4	.106	.212*
3	.042	3	.127	.127*
4	.051	2	.204	.102*
5	.041	1	.204	.041*
6	.016	0	.098	.000*
			.872	3.668

SECOND CASE
FLEET SIZE N = 4
MAINTENANCE FLOAT SIZE K = 4

i	P(i)	NO. OF UNITS IN FIELD	EXPECTED NO. OF UNITS IN SERVICE AND QUEUE	EXPECTED NO. OF UNITS IN FIELD
0	.493	4	0.0000	1.973*
1	.099	4	.099	.395 *
2	.039	4	.079	.158 *
3	.032	4	.095	.126 *
4	.051	4	.202	.202 *
5	.081	3	.404	.242 *
6	.097	2	.582	.194 *
7	.077	1	.543	.078 *
8	.031	0	.248	.000 *
			2.252	3.368

*These are intermediate values in a calculation, and do not in themselves reflect expected values.

INCREASE RELIABILITY OF OPERATIONAL SYSTEMS (IROS) **INDEX SERIAL NUMBER — 1074**

Earl D. Hendricks
IROS Program Manager
Sacramento Air Materiel Area (SMAMA/MMER)
McClellan Air Force Base, Calif. 95652

and

Alan K. Olsen
Chief, Quality Management Division
San Antonio Air Materiel Area (SAAMA/DSQ)
Kelly Air Force Base, Texas 78241

Abstract

This paper describes the development and implementation of the Air Force Increase Reliability of Operational Systems (IROS) Program. Explanations of purpose and program direction, along with a sketch of the program history, are given. Activities of the Air Force Logistics Command's (AFLC) Reliability/IROS Working Group have resulted in the application of computerized math models which interface with Air Force data systems to establish resource allocation priorities in the areas of reliability, logistic support cost, operational availability, and system safety. Multiple discipline teams at both the working and management levels are utilized to assure effectiveness. Economic resource allocations and cost effective system modifications are achieved through the IROS concept as applied to operational systems.

Acknowledgements

Insufficient space is provided here to acknowledge the many people and organizations who have influenced the shape and direction of the IROS Program. However, at least the membership of the AFLC Reliability/IROS Working Group, which has the prime responsibility for development and implementation of the IROS Program, should be identified. Those who have this continuing responsibility are: at Hq AFLC, Wright-Patterson AFB, Ohio, Charles R. Feeley, Craig W. Gridley, Edward L. Parkinson, and Frank J. Ruther; at OCAMA, Tinker AFB, Oklahoma, C. B. Shepherd, Jr.; at OOAMA, Hill AFB, Utah, Angelo Pappas; at SAAMA, Kelly AFB, Texas, Guy A. Morgan and Ben L. Williams; at SMAMA, McClellan AFB, California, Jack F. Bussio and John J. Fabish; and at WRAMA, Robins AFB, Georgia, William L. Richardson. Computer planning and programming support is provided by Elliot S. Leonard, Hq AFLC, and Eleanor L. Puckett, SMAMA.

Introduction

Presentation of a description of the IROS Program is given here to familiarize the government and industry professionals who are the key to its total success and to solicit their aid and support for it. Support of this type will eliminate redundancy and accomplish a step toward presentation of a unified and consistent approach by the profession to work toward common interests within government and industry defense organizations.

Background

In 1965 the late General J. P. Gerrity, then on the Air Force Staff, initiated the IROS Program. It was a sincere belief on his part that many of our operational defense systems could be improved in the area of reliability by incorporating the new products of an advancing technology or changing existing procedures, thereby reducing the tremendous cost of logistic support to these systems for their remaining life. General Gerrity was later assigned as Commander, Air Force Logistics Command, where the IROS Program proceeded toward development and implementation. There was an immediate review of all major systems, the intention being to identify those portions of the system which could be improved and would result in an overall economic benefit. Concurrently, steps were taken to develop computerized processes which would utilize operational and support data to provide a continual objective monitoring and priority establishing mechanism from which to select portions of systems to further evaluate for their potential as IROS candidates for improvement. General Gerrity's untimely death in the summer of 1968 resulted in a curtailment of the system review activity, although some items had been identified for improvement and were pursued to completion. Meanwhile, development of the computerized models continued at the same rate until initial implementation began in 1970. Considerable confusion occurred at this time and subsequently, by those who thought of the application of computerized models as a duplication of what had been done earlier; by those who were not aware of the IROS Program scope, and thought one particular computer model to be the entire IROS Program; and by those who had been involved in some Air Force or Department of Defense directed special IROS studies and saw no relationship with what they had done. To overcome this situation, it has been proposed that the program name be changed to System Effectiveness. This has been done in part in some organizations, but no matter what it is called, the program is continuing to complete implementation of the computerized models and application methods required to establish resource allocation priorities. This will broaden the information context out of which configuration of procedural change decisions are made. This enables decisions to be made on a system effectiveness versus cost basis. This will significantly impact the management decision process in the Air Force.

Mathematical Models

Mathematical models have been developed to provide a relative ranking of the items and subsystems within a defense system. These rankings use the five digit Work Unit Code (WUC) as a means of item identification. Rankings are in descending order of priority of the parameter estimated by the particular model. Models interface with field data and are used for Reliability Engineering evaluation and other system effectiveness evaluations leading to Configuration

Control Board (CCB) decisions on Air Force Class IV Modifications. A question might arise as to the validity of these products because of concern directed to accuracy of the field data, which is the operational phase input source. The answer to this question is covered by: (1) there are no particular reasons for maintenance personnel to bias the data by recording events which do not occur, or (2) if the data are biased, investigation will reveal the cause of this bias. In either case, the relative ranking provides a valid priority for investigation. The math models, functions, and references are as follows:

1. Mission Success Reliability, Reliability Engineering Evaluation, 1 and 2

2. Logistics Support Cost Ranking, Class IVC Modification (Logistics), 3

3. System Availability, Class IVB Modifications (Mission Essential), 4

4. System Safety, Class IVA Modification (Safety), 5

Mission Success Reliability

The Mission Success Reliability Math Model implemented in 1970 is a computerized general probability prediction model which processes apportioned, field or design data. The model is a fault tree, success path generating type which will accept any series/parallel design configuration as a one-time input maintained current with the configuration changes. Coding the configuration input, Figures 1 and 2, requires an understanding of reliability block diagrams, Figures 3, 4, and 5, but the model eliminates the necessity to write any equations relating the probabilities to the configuration. Subprograms accommodate total, partial, and standby redundancies singly or in nested combinations as well as equivalent blocks or crossovers. The output provides a multi-level ranking, Figures 6 and 7, of the entire system and may be used to study the complete mission or phase thereof. Use of this model during the new system full-scale development and production phases will provide uniformity and a better understanding of the results of the predictive analysis, and will eliminate the additional configuration input requirement if we are to use the model as part of the IROS Program during the deployment phase. This model has the characteristics necessary to fulfill the requirements of the reliability model and prediction portion of Contract Data Item R-3535/R-103, Reliability and Maintainability Allocations, Assessments and Analysis Report. The model also has the technical characteristics necessary to fulfill the total requirements of Contract Data Item R-3541/R-109, Computer-Programmed Mathematical Model for Reliability. A complete description of the Mission Success Reliability Math Model and its use may be found in references (1) & (2).

Logistics Support Cost Ranking

The Logistics Support Cost (LSC) Ranking Model development began at San Antonio Air Materiel Area (SAAMA) in 1968. Prototype model runs were made in 1969 and early 1970 and the LSC was implemented shortly thereafter in 1971. Continued refinement of the LSC has ensued.

Basically, the LSC is a cumulation of the most significant cost elements collected by component (designated in the Air Force by a Work Unit Code - WUC) on each defense system estimating the operational support cost for each WUC. The primary cost elements are base maintenance manhour costs, field shop costs, depot repair and overhaul costs, packing and shipping costs, replacement and condemnation costs and average base material costs for repair.

The primary output product ranks the WUCs by highest dollar value first and then in descending order by dollar value, figure 8. The proportionate share column shows the percent of the total defense system support cost attributed to that particular WUC. The three previous quarter costs and ranks illustrate the trend that is experienced by a WUC. The equivalent rate is a measure of the number of months required for a WUC's LSC to equal the WUC acquisition cost. Another product ranks the WUCs for the entire defense system in descending order of equivalent rate. Another product ranks the WUC in numerical sequence. The various products give access to data for the many different applications, figures 9, 10, and 11.

The LSC gives you a total defense system picture of support cost. In this sense, it points a finger at the areas which are the high resource consumers; hence, the items which need investigation for potential support cost reduction through some form of corrective action are identified. It has been observed that on most systems the top ten WUCs account for over 25 percent of the total costs. For the manager, this is relevant information in deciding whether or not to invest effort in problem isolation or investigation.

Similarly the equivalent rate sequence product points a finger at those items which have a high support cost in relation to their acquisition value. Many times, a simple change will allow dramatic reductions in these items.

System Availability

The System Availability Model (SAM) development began at SAAMA in 1970 and was implemented in 1972. The SAM provides a measure of defense system availability degradation due to each WUC. The four elements which are used as a basis for the calculation are Not Operationally Ready due to Supply (NORS), Not Operationally Ready due to Maintenance (NORM), ground aborts and flight aborts. The data is also available by aircraft serial, "tail" number.

This model has four products:

a. Rank Sequence by Work Unit Code, figure 12.

b. Rank Sequence by Aircraft Serial Number.

c. Work Unit Code Sequence.

d. Aircraft Serial Number Sequence.

The SAM gives you a total picture of availability degradation. In this way, it identifies those WUCs which need to be investigated for possible corrective action. For the manager, this model is a tool for resource allocation.

Flight Safety Prediction Technique

The Flight Safety Prediction Technique (FSPT) is a managerial and engineering tool for objective quantification of flight safety. The FSPT can be used to predict unsafe situations, to establish workload priorities and to evaluate proposed system modifications. The technique was developed by SAAMA on contract with ARINC Research Corporation beginning in September 1966.

The FSPT is a method of safety assessment which can be applied to any defense system in the Air Force inventory. FSPT implementation on all Air Force aircraft is in process. At the present time, three defense systems have operational models.

Two approaches were taken in generating safety indices to be used as pointers to potential problem areas within a defense system. One approach (Criticality Model) uses as input data from an AF-wide system; the other (State-Phase Model) is dependent on information from ADC Pilot Post-Flight Debriefings.

In the State-Phase Model, for each phase of an average flight, the pilot is assumed to be in one of three operational states: (1) Safe:No equipment/system malfunction symptom present; (2) Mode I Unsafe:Equipment malfunction present, but recovery or alternate mode of operation available; (3) Mode II Unsafe:Disaster imminent. The probabilities for being in each of the three states during each phase of an average flight due to each of the pilot-reported symptoms are calculated. Symptoms are reported in two digit codes such as "3C", the "3" designating Electrical Power/Landing-Warning and the "C" pinpointing the problem to AC/DC power failure. For summary purposes, various averages and rankings are performed on a monthly basis to furnish a quick portrayal of recurring problems as indicated by the malfunction symptoms which are experienced in flight.

The second method, utilizing maintenance data from the AFM 66-1 data system, is designed to handle WUC level inputs as a measure of fleet performance. The safety indices generated in this model, given on a WUC basis are functions of the severity of the loss of a given function and the probability that a given WUC will fail on an average flight. To access the severity of the loss of a given WUC, all functions which are dependent on this WUC must be determined. This is done via a functional diagram of the weapon system under study. Tabulated for each major function are (1) the equipment necessary for its performance, (2) operating modes of the equipment, and (3) all inputs required from other systems. For each flight safety related WUC, an associated "sensitivity" is obtained. This sensitivity is a measure of the significance of the loss of the WUC to the safe flight of the aircraft. When multiplied by the probability that the WUC will fail, a "criticality" for the given WUC is obtained, figure 13.

The FSPT gives you a picture of the WUCs which contributed to a safety hazard in the last reporting period. Again, this is a tool which the manager can apply to resource allocation for problem resolution.

IROS Groups

The IROS Program has developed along the lines of involving team action in the use of the math models to establish priorities throughout the evaluation and modification process. The team concept includes a working group and a management group for each major defense system, Figure 14. The working level groups are known at the various AMAs as the IROS Working Groups or the System Effectiveness Groups. The nucleus of these groups are composed of reliability engineers, equipment specialists, and data specialists. Their task is to provide continuous tracking, problem identification assistance, solution assistance, implementation assistance, and follow-up assessment. The working group is expanded to bring together particular specialists appropriate to study particular problems in order to arrive at proposed solutions and methods of implementation. The IROS Task Groups have broad representation (AMA, Using Command, AFSC, USAF Safety, Contractors, other) and meet as often as program activities dictate.

The math models are applied by the IROS Groups to assess the benefits and costs for each of the problem areas presented. The use of the LSC, SAM and FSPT models allow the IROS Group to address the question of how much performance could be gained in each of these three areas. These are the benefits in terms of outcome. The cost of the modification is then used to calculate the benefit/cost ratios for each performance category, Figure 15.

Configuration Control Board

The results of the analysis of the math models are integrated with information and recommendations from other sources for submittal to the Configuration Control Board. Each corrective action is considered with respect to, and assigned one of, three categories:

Class IVA Flight safety deficiency modification

Class IVB Mission essential modification

Class IVC Cost savings modification

Therefore, each proposed modification has an objective, evaluation, and priority established in the performance and cost areas for the CCB to review and take final action.

References

1 AFLCM 66-18, Programming and Technical Processes, Chapter 18, Mission Success Reliability Mathematical Model.

2 AFLCM 171-229, Weapon System Reliability Mathematical Model Program (K051).

3 AFLCM 66-18, Programming and Technical Processes, Chapter 17, Logistics Support Cost Ranking.

4 ibid, Chapter to be published, System Availability.

5 ibid, Chapter to be published, System Safety.

6 AFR 80-5, Reliability and Maintainability Programs for Systems, Subsystems, Equipment, and Munitions.

KØ51 RELIABILITY BLOCK DIAGRAM PREPARATION — PAGE 1 OF 1 PAGES

PREPARED BY: John Doe — ORGANIZATION: OOAMA/MMER

SECTION A – TITLE BLOCK

1. DOCUMENT IDENTIFIER (1 – 3)	2. MDS MISSION (4 – 6)	MDS DESIGN (7 – 9)	MDS SERIES (10)	3. DIAGRAM NUMBER FUNCTION NUMBER (11 – 12)	SUB-SYSTEM NO. (13 – 14)	BLOCK NUMBER (15 – 18)	4. MISSION APPLICATION (19 – 30) 1	2	3	4	5	6	7	8	9	10	11	12	5. DIAGRAM TITLE (31 – 54)	6. JULIAN DATE (75 – 79)	7. CHG CODE (80)
MK1	F	4	C	1 5	Ø 3	Ø Ø	1	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Q R C - 1 6 Ø A - Z, E C M P O D	7 2 2 2 4	A

SECTION B – BLOCK CARDS

CARD 1

8. DOC IDENT (1–3)	9. BLK NO. (15–16)	10. WORK UNIT CODE/BLOCK REF NO. (19 – 24)	11. EXP (29–30)	12. BLOCK TITLE (31 – 54)	13. EQUIVALENT BLOCK NUMBERS (55 – 70) 1	2	3	4	5	6	7	8	14. CC (80)
MK2	6 4	7 6 2 L D		P O W E R S U P P L Y									A
	6 5	7 6 2 L F		R E C E I V E R									A

CARD 2

15. INPUT BLOCK NUMBERS (19 – 38) 1	2	3	4	5	6	7	8	9	10	16. OUTPUT BLOCK NUMBERS (39 – 58) 1	2	3	4	5	6	7	8	9	10	CC (80)
6 3										6 6										A
6 3										6 6										A

SECTION C – REDUNDANCY CARDS

1. DOCUMENT IDENT (1 – 3)	2. BLOCK NO. (15 – 16)	3. CARD NO. (17 – 18)	4. MINIMUM REQUIRED (19 – 20)	5. MAXIMUM AVAILABLE (21 – 22)	6. REDUNDANT COMPONENTS WUC/BLOCK REF NO. (23 – 28)	DELIMITER (29)	WUC/BLOCK REF NO. (30 – 35)	DELIMITER (36)	WUC/BLOCK REF NO. (37 – 42)	DELIMITER (43)	WUC/BLOCK REF NO. (44 – 49)	DELIMITER (50)	WUC/BLOCK REF NO. (51 – 56)	DELIMITER (57)	7. CHG CODE (80)
MK3	2 3	Ø 1	Ø 1	Ø 2	7 6 2 L D	*									
	2 3	Ø 2	Ø 1	Ø 2	7 6 2 L F	*									

Figure 1 K051 Reliability Block Diagram Preparation

KØ51 COMPONENT DATA FILE MAINTENANCE — PAGE 1 OF 1 PAGES

PREPARED BY: John Doe — ORGANIZATION: SMAMA/MMER

SECTION A – MISSION TIME CARD

1. DOC IDENT (1 – 3)	2. MDS MISSION (4 – 6)	DESIGN (7 – 9)	SERIES (10)	3. MISSION/MISSION PHASE TIMES 1 (21 – 24)	2 (25 – 28)	3 (29 – 32)	4 (33 – 36)	5 (37 – 40)	6 (41 – 44)	7 (45 – 48)	8 (49 – 52)	9 (53 – 56)	10 (57 – 60)	11 (61 – 64)	12 (65 – 68)	4. PRIME AMA (73 – 74)	5. JULIAN DATE (75 – 79)	6. CHANGE CODE (80)
MD1	F	1 Ø Ø	D	Ø Ø 1 5												S M	7 2 Ø 4 2	A

SECTION B – MAR TABLE CARDS

CARD 1

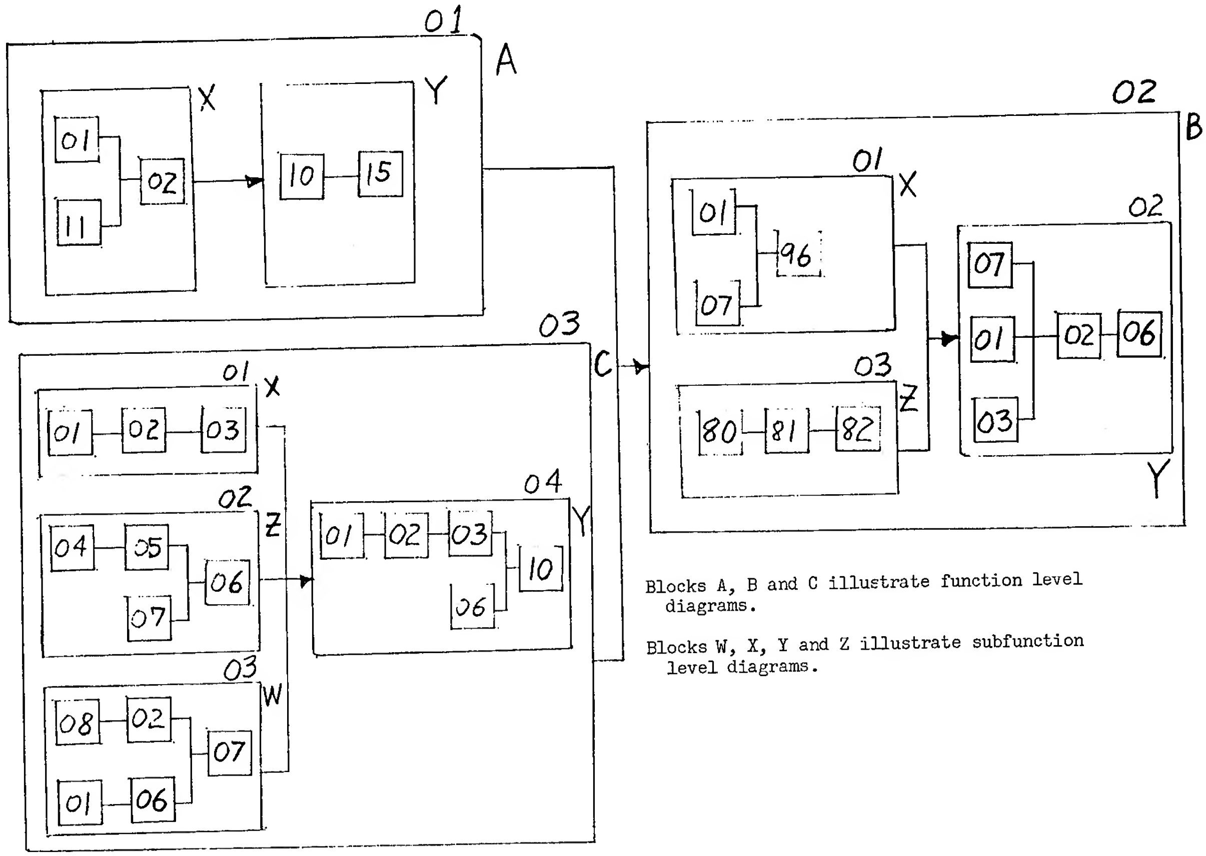
1. DOC IDENT (1 – 3)	2. MDS MISSION (4 – 6)	DESIGN (7 – 9)	SERIES (10)	3. WORK UNIT CODE (11 – 16)	4. CARD NO. (18)	5. MAR TYPE CODE (19)	6. MINIMUM ACCEPTABLE RELIABILITY (20 – 28)	7. QUANTITY PER APPLICATION (29 – 33)	8. CHG CODE (80)
MF1	F	1 Ø Ø	D	6 3 1 1 D	1	H	Ø Ø Ø Ø Ø Ø 8 8 5	Ø Ø Ø Ø 1	A
					1				
					1				
					1				
					1				
					1				
					1				
					1				
					1				
					1				
					1				
					1				
					1				
					1				
					1				

CARD 2

9. CARD NO. (19)	10. OP TIME CODE (20)	11. MISSION/PHASE OPERATE TIME OR CYCLES 1 (21 – 24)	2 (25 – 28)	3 (29 – 32)	4 (33 – 36)	5 (37 – 40)	6 (41 – 44)	7 (45 – 48)	8 (49 – 52)	9 (53 – 56)	10 (57 – 60)	11 (61 – 64)	12 (65 – 68)	12. CHG CODE (80)
2	H	Ø Ø 1 5												A
2														
2														
2														
2														
2														
2														
2														
2														
2														
2														
2														
2														
2														
2														

AFLC FORM JUN 70 122 FC 4400 AFLC–WPAFB–JUN 70 10M

Figure 2 K051 Component Data File Maintenance



Blocks A, B and C illustrate function level diagrams.

Blocks W, X, Y and Z illustrate subfunction level diagrams.

Figure 3 Weapons System Level Block Diagram

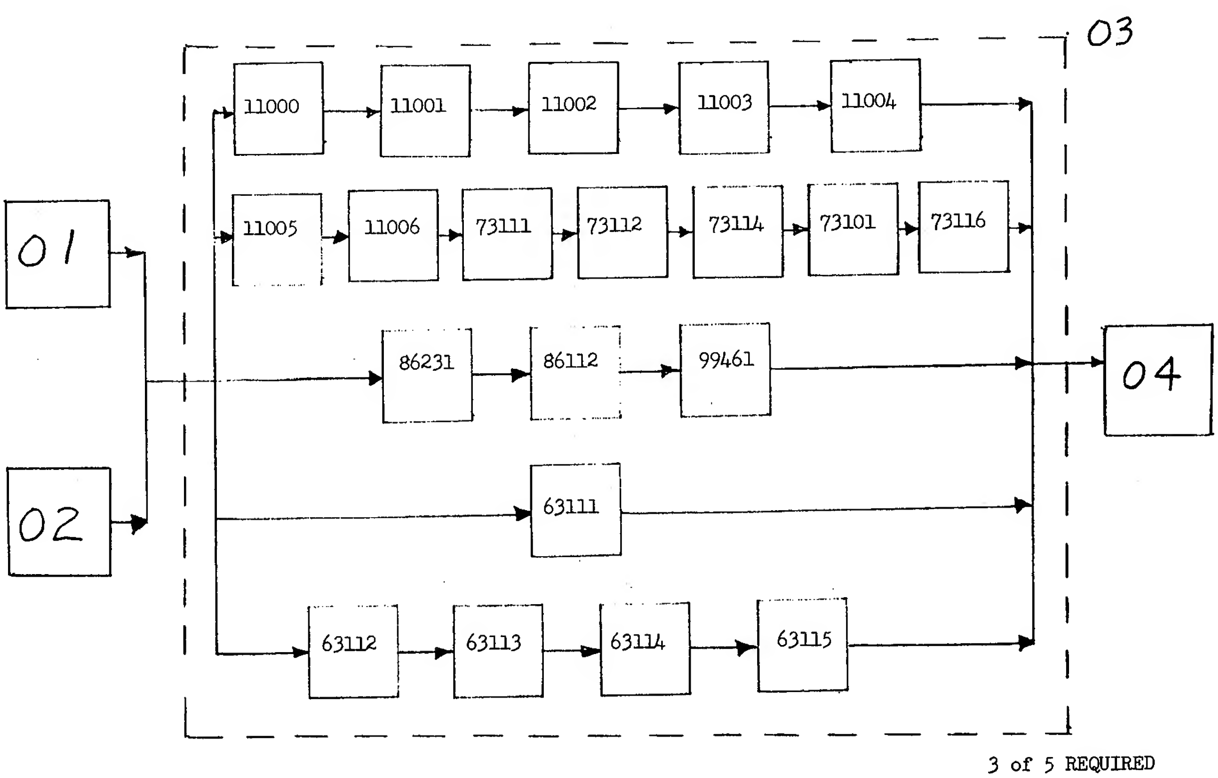


Figure 4 Diagram Containing a Partially Redundant System

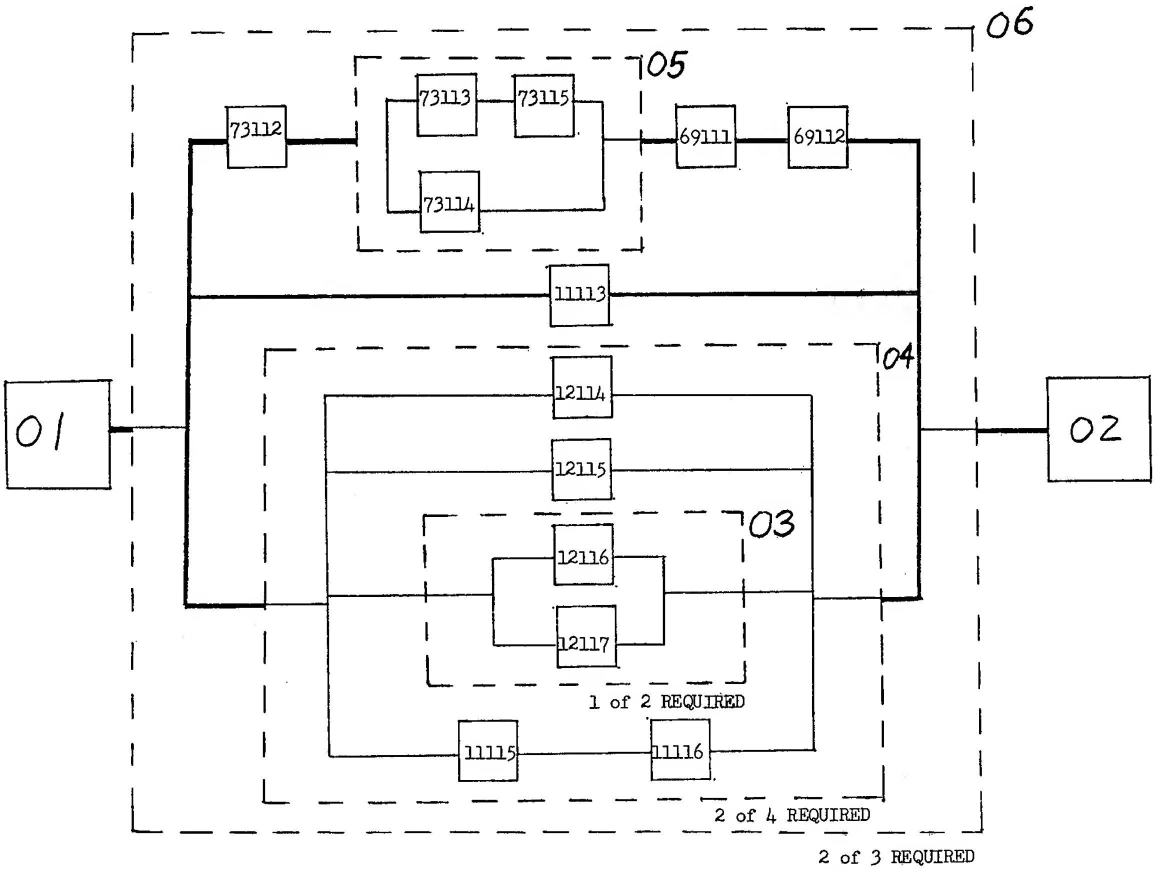


Figure 5 Nested Total and Switching Redundancies

REQUESTOR - SMAMA-MMER STUDY PROBABILITIES RCS-LOG- K051.RL1L 68 SEP 27 PAGE 0002

STUDY 927 FINAL CONFIGURATION TEST RF111F MISSION 04

A. SYSTEM PROBABILITY OF SUCCESS 0.79999999

	DIAGRAM F SF BK	WUC	PROB OF SUC	DESCRIPTION
B. LOWEST FUNCTION PROBABILITY	16 00 00		.72236791	PROPULSION SYSTEM
LOWEST INCLUDED SUBFUNCTION PROBABILITY	16 32 00		.72146977	ENGINE INSTRUMENTS
LOWEST INCLUDED COMPONENT PROBABILITY	16 32 11	12345	.69998766	EXHAUST TEMPERATURE IND
NEXT LOWEST FUNCTION PROBABILITY	18 00 00		.83165412	WEAPON CONTROL
LOWEST INCLUDED SUBFUNCTION PROBABILITY	18 31 00		.80000000	EXTERNAL STORES
LOWEST INCLUDED COMPONENT PROBABILITY	18 31 69	13246	.91111111	GROUND LOCK-OUT RELAY
NEXT LOWEST FUNCTION PROBABILITY	13 00 00		.84112234	INFORMATION DISPLAY
LOWEST INCLUDED SUBFUNCTION PROBABILITY	13 26 00		.88765432	ENGINE INSTRUMENTS
LOWEST INCLUDED COMPONENT PROBABILITY	13 26 01	51311	.90000000	TACHOMETER GENERATOR
NEXT LOWEST FUNCTION PROBABILITY	01 00 00		.93411326	FIRE CONTROL
LOWEST INCLUDED SUBFUNCTION PROBABILITY	01 10 00		.87932161	GUN RADAR
LOWEST INCLUDED COMPONENT PROBABILITY	01 10 17	97965	.90998791	RADAR POWER DISTRIBUTION

A - SYSTEM PROBABILITY OF SUCCESS IS GIVEN ONLY WHEN THE ENTIRE SYSTEM IS STUDIED. OTHERWISE, A STUDY PROBABILITY OF SUCCESS WILL BE GIVEN COVERING THE RANGE OF THE STUDY. IF ALL SUBFUNCTIONS WITHIN A FUNCTION ARE STUDIED THIS FIGURE WILL BE A FUNCTION PROBABILITY OF SUCCESS. IF ONLY A PORTION OF THE SUBFUNCTIONS ARE STUDIED, THE PROBABILITY IS ONLY FOR THE RANGE OF THE STUDY

B - FUNCTION PROBABILITIES ARE OMITTED WHEN STUDIES ARE ON THE SUBFUNCTION LEVEL.

Figure 6 Study Probabilities

REQUESTOR - SMAMA-~~NER~~ MMER   BLOCK PROBABILITIES   RCS-LOG-   K051.RL2L   68 SEP 10   PAGE 0001

STUDY 047   TRIAL DESIGN   RF111F   MISSION 04

SUBFUNCTION 37 01 00   CONTAINMENT   PROBABILITY OF SUBFUNCTION SUCCESS 0.93711460

BLOCK	PROB OF SUC	WUC	DESCRIPTION	MAR	FIELD DATA	DESIGN DATA
01	1.00000000	79111	FUSELAGE FWD. SECTION	990000.0 MTBF		1.00000000 PROB
02	0.93400000	79112	FUSELAGE MIDDLE SECTION	7000000.0 MTBF	9734000.0 MTBF	
03	0.87000000	79113	FUSELAGE AFT SECTION	670000.0 MTBF	17770000.0 MTBF	

SUBFUNCTION 37 02 00   PROPULSION   PROBABILITY OF SUBFUNCTION SUCCESS 0.64073216

BLOCK	PROB OF SUC	WUC	DESCRIPTION	MAR	FIELD DATA	DESIGN DATA
01	0.90000000	63111	ENGINE BASIC	7367.0 MTBF		
02	0.87650000	63112	NOSE ACCESSORY	41007.0 MTBF		
03	0.937A0000	63113	MAIN ACCESSORY	7113700.0 MTBF		
04	1.00000000	63114	ENGINE OIL	6310072.0 MTBF		1.00000000 PROB
05	0.73100700	63115	IGNITION AND START	37611.0 MTBF		.73100700 PROB

SUBFUNCTION 37 03 00   AUXILIARY POWER   PROBABILITY OF SUBFUNCTION SUCCESS 0.73461000

BLOCK	PROB OF SUC	WUC	DESCRIPTION	MAR	FIELD DATA	DESIGN DATA
07	0.93700000	74333	DC ELECTRICAL	63700.1 MTBF	463210.0 MTBF	
09	0.89773700	74334	AC ELECTRICAL	7632.0 MTBF	100000000 MCBF	
11	0.82113700	74335	UTILITY HYDRAULIC P.S.	96374 MCBF	73641.0 MTBF	
13	0.41637000	74336	WIRING AND CONNECTORS	99375668 MCBF	11111000 MCBF	

FUNCTION 37 99 00   AUXILIARY POWER PLANT   PROBABILITY OF FUNCTION SUCCESS 0.56341100

BLOCK	PROB OF SUC	WUC	DESCRIPTION	MAR	FIELD DATA	D DESIGN DATA
01	0.93711460	37 01 00	CONTAINMENT	246731.0 MTBF		
02	0.64073216	37 02 00	PROPULSION	71100.0 MTBF		
03	0.73461000	37 03 00	AUXILIARY POWER	630000.0 MTBF		

Figure 7 Block Possibilities

WEAPON SYSTEM F004C OOAMA   LOGISTIC SUPPORT COST RANKING   K051.PN1L   PAGE 1
AFM 65-110/66-1 DATA AS OF 72 JUN   SELECTED ITEMS   DATE PROCESSED 72 JUL 25

			AVERAGE MONTHLY VALUES									
		PROP	CURRENT QTR		1ST PREV QTR		2ND PREV QTR		3RD PREV QTR		EQUIV	RATE
WUC	NOUN	SHARE	RANK	LSC	RANK	LSC	RANK	LSC	RANK	LSC	RANK	MONTHS
71E60	PLATFORM GYRO S	4.391	1	$74,183	1	$75,122	P 126	$2,270	P 225	$1,212	148	138.1
23720	CONST SPEED DRI	3.214	2	$54,300	2	$62,067	1	$52,842	2	$56,716	101	87.4
23200	ENGINE &QEC	2.992	3	$50,538	3	$59,898	2	$42,801	1	$69,767		
42210	GEN 3 PHASE 400	2.244	4	$37,905	6	$32,161	4	$32,773	5	$25,438	22	22.0
1226F	BUCKET SEAT	2.008	5	$33,920	4	$39,223	3	$40,164	3	$35,764	1	0.2
71LA0	REC-TRANS RT-54	1.874	6	$31,655	5	$36,405	5	$30,612	4	$28,107	12	17.7
75130	AERO 7-A	1.681	7	$28,395	7	$24,582	6	$25,545	9	$18,136	161	148.1
111AA	RADOME	1.570	8	$26,521	15	$14,426	19	$12,187	20	$10,877	70	63.9
71E20	COMPUTER NAVIGA	1.480	9	$25,004	10	$19,927	7	$21,046	8	$19,449	216	220.8
23730	CONSTANT SPD DR	1.408	10	$23,791	9	$21,645	11	$15,515	35	$6,595	204	204.5
73160	DISPLACEMENT GY	1.266	11	$21,386	P 352	$717	P 225	$1,246	P 306	$823	91	76.5
71B10	CONTROL COMPUTE	1.184	12	$20,000	14	$15,031	10	$18,612	6	$21,925	197	195.6
74100	RADAR SET APQ-1	1.159	13	$19,579	12	$18,969	9	$18,722	7	$19,782		
74110	SYNCH ELEC SN-3	0.959	14	$16,203	31	$8,856	31	$8,142	45	$5,251		
74120	CONT POWR SUP C	0.948	15	$16,010	11	$19,155	8	$19,160	10	$17,389	141	125.8

NOTE -- THIS REPORT IS ON MICROFILM AND PRINTOUT.
IT REFLECTS AT LEAST THE TOP 150 ITEMS AND ANY ITEMS WITH AN LSC GREATER THAN $100.

Figure 8 Logistic Support Cost Sequence

WEAPON SYSTEM F004C OOAMA — LOGISTIC SUPPORT COST RANKING — K051.PN3L PAGE 1
AFM 65-110/66-1 DATA AS OF 72 JUN — ·WORK UNIT CODE STATUS — DATE PROCESSED 72 JUL 25

	AVERAGE MONTHLY VALUES			
FLEET INFORMATION	CURRENT QTR	1ST PREV QTR	2ND PREV QTR	3RD PREV QTR
LOGISTIC SUPPORT COST	$167,886	$136,931	$152,629	$184,743
INVENTORY	331	331	345	352
OPERATING HOURS	4614	6533	6690	7330
LANDINGS	5551	5326	5065	6267
23 ENGINE INFORMATION				
AIRCRAFT INVENTORY	331	331	345	352
OPERATING HOURS	9228	13066	13380	14660

			AVERAGE MONTHLY VALUES									
		PROP	CURRENT QTR		1ST PREV QTR		2ND PREV QTR		3RD PREV QTR		EQUIV	RATE
WUC	NOUN	SHARE	RANK	LSC	RANK	LSC	RANK	LSC	RANK	LSC	RANK	MONTHS
1226B	REEL INERTIA	0.047	P 360	$799	P 443	$548	P 261	$1061	P 184	$1444	P 371	706.2
1226C	STRAP SHOULDER	0.000	1903	$4	1471	$31						
1226D	PLATE PARA MOUN	0.000			1784	$8	1566	$23	1384	$33		
1226F	BUCKET SEAT	2.008	5	$33920	4	$39223	3	$40164	3	$35764	1	0.2
1226K	PARACHUT DROGU	0.038	P 411	$638	338	$742	294	$292	294	$851	P 282	342.2
122XX				$35361		$40552		$42177		$38092		
1232D	STRUCTURE	0.036	424	$603	549	$401	1219	$69	1077	$86		
12320	FWD CANOPY ASY	0.101	197	$1698	49	$5521	P 78	$3412	P 54	$4358	311	419.2
123XX				$2301		$3481		$4444				
12XXX		2.492		$37662		$46474		$45658		$42536		
1325D	TIRES MLG L-H	0.262	72	$4434	30	$8916	1428	$38	26	$7814	6	10.2

NOTE -- THIS REPORT IS ON MICROFILM AND PRINTOUT.

Figure 9 Work Unit Code Sequence

WEAPON SYSTEM F004C OOAMA — LOGISTIC SUPPORT COST BREAKDOWN — K051.PN4L PAGE 1
AFM 65-110/66-1 DATA AS OF 72 JUN — CURRENT QUARTER COMPUTATION — DATE PROCESSED 72 JUL 25

		QUARTERLY VALUES					
WUC	NOUN	FIELD MAINT	SPEC REPAIR COST	PACK/SHIP COST	CONDEMNATION COST	BASE MATERIAL COST	TOTAL QTR LSC
1226B	REEL INERTIA LO	$2,387	P $0	$10	$0	$0	P $2,397
1226C	STRAP SHOULDER	$12	$0	$0	$0	$0	$12
1226F	BUCKET SEAT	$101,760	$0	$0	$0	$0	$101,760
1226K	PARACHUT DROGU	$1,914	$0	$0	P $0	$0	P $1,914
122XX		$106,073	$0	$10	$0	$0	$106,083
1232D	STRUCTURE	$1,234	$329	$42	$0	$204	$1,809
12320	FWD CANOPY ASY	$2,255	$605	$78	$2,117	$ 39	$5,094
123XX		$3,489	$934	$120	$2,117	$243	$6,903
12XXX		$109,562	$934	$130	$2,117	$243	$112,986
1325D	TIRES MLG L-H		7,492	$932	$865	$329	$13,302

NOTE -- THIS REPORT IS ONLY ON MICROFILM.

Figure 10 Logistic Support Cost Breakdown

WEAPON SYSTEM F004C OOAMA LOGISTIC SUPPORT COST FILE MAINTENANCE REGISTER K051.PN7L PAGE 1

CHANGED DATA ELEMENTS ARE DENOTED BY C <COMPUTER> OR M <MANUAL> DATE PROCESSED 72 JUL 25

FLEET INVENTORY 331 PERCENT CONUS 50

WUC	ACC	FSN	PRIME AMA	NOUN	ERRC	UNIT PRICE	WEIGHT	QPA	SRA REPAIR	% DEPOT CONDEMN	BASE MAT COST	OLD WUC
1226B	C G	13771187468	D	REEL ASSEMBLY,I C	T C	$852.40	C 6		C $72	-- C	$0	12286
	G	C 13771278360	D	REEL INERTIA LO	T C	$852.40	6		------	--	$0	12286
		13771288380	***	CHANGE TO 13771278360 ***								
	F	9999999999999		REEL INERTIA LO	-	$852.40		2				12286
1226C	G	29256842355FS	***	DELETED NO FIELD DATA HISTORY ***								
	C Y	9999999999999		M STRAP SHOULDER	-	----------		M 2				
1226D	Y	9999999999999		PLATE PARA MOUN	T	----------		2				
1226F	M G	M 13778718153BF	M E	---------------	M T C	$10.21	C 1		------	C 00	$0	
REJECTED M2E		F004C1226F 29256842222		D	00045000			72187C	*** ILLOGICAL F/M ACTION			
1226F	F	9999999999999		BUCKET SEAT	-	$10.21		2				
1226K	G	16708903615BF	E	DROGUE CHUTE AS	X	$329.60	6		------	00	$0	
	F	9999999999999		PARACHUTE DROGU	X	$329.00		2				
1232D	Y	9999999999999		STRUCTURE	-	$1050.00		1				
12320	G	15601048136BF	E	CANOPY,MOVABLE	T	$2183.00	256		C $1317	03	$0	
	G	15607406974BF	E	CANOPY,MOVABLE	T	$2117.00	250		C $1307	C 02	$50	
	F	9999999999999		FWD CANOPY ASSY	-	$2150.00		1				
1321F	X	9999999999999		VALVE,RESTRICTO	N	$1.00		1				
133XX			***	DELETED ***								
14XXX			***	DELETED ***								
15ABC			***	DELETED ***								
2321B	G	28406903727	M J	DUCT,TRANSITION M	X	$891.30	40		$399	00	$15	
REJECTED M2E		F004C2321B 28406903728		E		45 01	240	72287A				
					.************		*****					
2321B	G	49205944703	D	ARM,E,TENSION	U	$87.01	6		------	00	$0	
	A	9999999999999		INLET GUIDE VAN	-	----------		2				

NOTE--THIS REPORT IS ONLY ON MICROFILM.

Figure 11 Logistic Support Cost File Maintenance Register

S Y S T E M A V A I L A B I L I T Y R A N K I N G

TOP 20 WORK UNIT CODES

RANK	REFERENCE REPORT ITEM/DATE	WUC	NOUN	% OF TOTAL DEGRADATION
1	5 NOV 71	23Z00	ENGINE/QEC KIT	05.353
2		23NDC	THERMOCOUPLE PROBE	03.747
3	1 AUG 71	23GAB	FUEL CONTROL	03.650
4		14000	FLIGHT CONTROLS	03.083
5		23000	TURBOJET POWER PLT	03.038
6		23MDD	THERMO HARNESS	02.278
7		4622A	WING VENT VALVE NT	02.041
8		42123	MAIN SOLID STATE	01.682
9	9 FEB 72	47211	OXYGEN REGULATOR	01.390
10		46211	BOOST PUMP RG1110Ø	01.354
11		11515	COMPRESSION LINK	01.283
12	2 AUG 71	63115	RECEIVER XMITTER	01.270
13		51132	ATTITUDE IND J8	01.044
14		4112A	COOLING TURBINE	01.031
15		47112	OXYGEN FILLER VALVE	01.025
16		49112	OVERHEAT DET CABLE	01.018
17	8 NOV 71	51133	ATTITUDE IND MM3	00.966
18		13212	NOSE STRUT ASSEMBLY	00.951
19		46000	FUEL SYSTEM	00.949
20		0411J	OPR READY INSP	00.803

Figure 12 System Availability Ranking

S Y S T E M   S A F E T Y   R A N K I N G

TOP 20 WORK UNIT CODES

RANK	REFERENCE REPORT ITEM/DATE		WUC	NOUN	CRITICALITY X10**(-3)
1			46211	BOOST PUMP RG11100	.729
2			42123	MAIN SOLID STATE	.290
3			13212	NOSE STRUT ASSEMBLY	.233
4	2	AUG 71	63115	RECEIVER XMITTER	.147
5			13722	NOSE TIRE	.133
6	1	AUG 71	23GAB	FUEL CONTROL	.119
7			46212	PROPORTIONER	.106
[illegible]			13712	NOSE WHEEL	.100
9			42211	VOLTAGE REGULATOR	.092
10			13142	DOWN LOCK SWITCH	.068
11			1313B	ACTUATING CYL	.067
12	4	AUG 71	13112	MAIN STRUT ASSEMBLY	.067
13			13211	NOSE GEAR ASSEMBLY	.067
14			13721	MAIN TIRE	.067
15	8	NOV 71	51133	ATTITUDE IND MM3	.065
16	3	AUG 71	13411	BRAKE ASSEMBLY	.064
17			71113	INSTRUMENTAL UNIT	.057
18			71115	RECEIVER	.052
19			51132	ATTITUDE IND J8	.049
20			46318	T-HANDLE	.040

Figure 13 System Safety Ranking

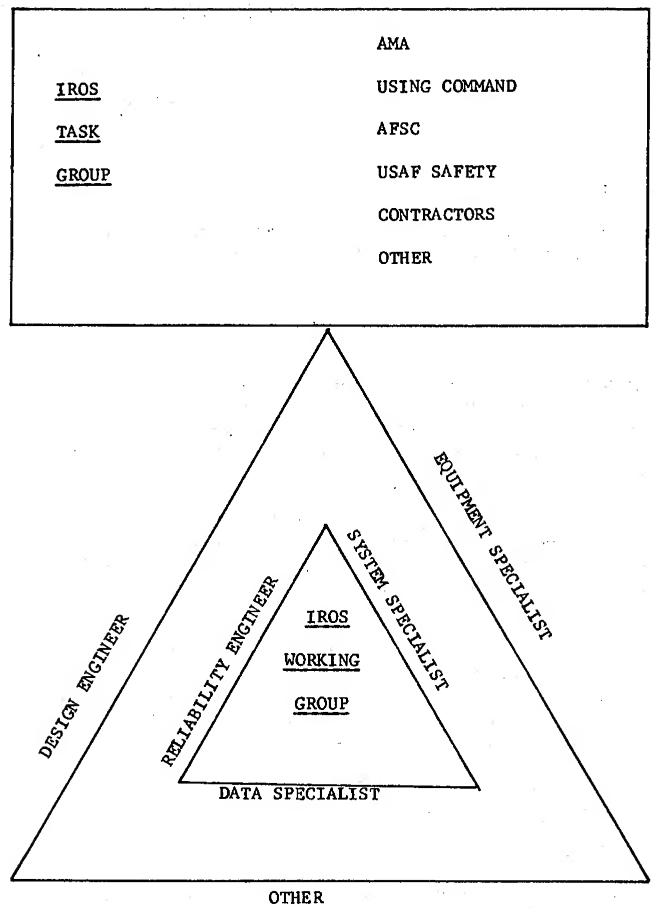


Figure 14

CONFIGURATION CONTROL BOARD SYSTEM EFFECTIVENESS SUMMARY — DATE: 18 Apr 72

TITLE: INVERTER Switch (Relay Control) — MIP NUMBER: SANAA 72-0011 — CONTROL NUMBER: O-153

MODIFICATION CLASS: ☒ IVA (Safety) 1684 ☐ IVB (Mission Essential) ☐ IVC (Logistics)

SYSTEMS/EQUIPMENT AFFECTED: T-37 — WUC 42233

EXISTING SYSTEM	PROPOSED SYSTEM (PREDICTED VALUES)	CHANGE	BENEFIT/DOLLAR
LOGISTIC SUPPORT COST $956/MO RANK 183	LOGISTIC SUPPORT COST $95/MO RANK 526	$861/MO.	$ RETURN/$ SPENT 0.144
AVAILABILITY (% DEGRADATION) 0.02 RANK 136	% DEGRADATION 0.002 RANK 456	0.018	AIRCRAFT EQUIVALENT/$10K 0.022 AIRCRAFT EQUIVALENT .16
SAFETY (CRITICALITY X 10⁶) 40 RANK 20	CRITICALITY X 10⁶ 4 RANK 75	36	CHANGE IN CRITICALITY X 10⁶/$10K 4.86

- AMORTIZATION TIME 6.91 YEARS
- MODIFICATION COST $71,400
- SAVINGS $10,331 PER year

AFLC FORM MAR 72 48A

Figure 15 Cost/Benefit Calculations

LOGISTICS EFFECTS ON OPERATIONAL AVAILABILITY

by Herbert G. Jacks
Staff Engineer
Reliability Assurance Department
Librascope Division of
The Singer Company
Glendale, California

Abstract:

The commonly used equipment Availability equation considers only mean time between failures (M) and mean time to repair (μ). It may be written as

$$A = \frac{M}{M+\mu} \tag{1}$$

This equation is rigorous only if the deployment period is infinite and repair capability is unlimited, but gives useful results for situations in which equipment deployment period is fairly long compared to mean time between failures, and repair capability adequate to correct most failures.

Electronic equipment installed at advanced military bases or on small ships (such as destroyers and submarines, on which both repair capability and deployment period are limited) is an important example of equipment installations for which use of the conventional Availability equation is inappropriate.

In order that computed Availability be a valid measure of the probability that the equipment will be operable at a random point in time, it is first necessary to determine the probability of operation at each point in time, and then to compute the average value of that probability function over the entire deployment period (T).

The probability that the equipment will be operable at a particular point in time (t) depends on the probability that the equipment has not failed and the probability that, if the equipment has suffered a failure, repair has been completed prior to time (t).

Where repair capability is limited by the supply of spares to (n) failures, Availability may be computed by use of the equation

$$A_n(T) = \frac{1}{T}\int_0^T e^{-\frac{t}{M}} + \frac{M}{M+\mu} \sum_{i=1}^{n} \frac{\left(\frac{t}{M}\right)^i e^{-\frac{t}{M}}}{i!} \, dt$$

INTRODUCTION

It is common practice to make Availability analyses of systems and equipment using the equation

$$A = \frac{M}{M+\mu} \tag{1}$$

where M is the mean time between failures and μ is the mean time to repair the equipment.

While (1) is convenient, reasonably accurate, and useful for many analyses, it is rigorous only if on-site repair capability is unlimited and the deployment period is infinite.

Failure to give proper consideration to the limitations of equation (1) can result in computed Availability values which are meaningless as a measure of equipment operational value to the user.

LOGISTICS AND SUPPORT DELAYS

Limitations on maintenance capability fall into two general classes:

- Delays in performing maintenance because of administrative delays and time spent hunting for spare parts which are at the operating site.
- Inability to make a repair due to lack of spare parts, special tools, or adequately skilled personnel.

The delays resulting from maintenance personnel being busy repairing other equipment or hunting through spare parts stocks for the required part are important; however, the delay in maintenance which results from unavailability of the required spare parts at the operating site is often the most important factor in operational availability.

In the particular case of complex electronic systems installed on destroyers and submarines, the relatively small percentage of not-repairable-on-board failures commonly results in considerably more equipment down time than that associated with all repairable failures.

EFFECT OF FINITE DEPLOYMENT PERIOD

A completely different source of errors in estimating equipment operational availability is that which results from the implicit assumption in equation (1) that the equipment being evaluated has been operating since the dawn of time. Errors resulting from this assumption are ordinarily negligible for low MTBF fixed station equipments (radars, large computers, etc.) which may fail and be repaired dozens of times during their periods of deployment. However, the errors resulting from this "steady state" assumption are far from negligible for a wide variety of operational situations in which the number of failures during a deployment period is small or zero.

It is the intent of this paper to indicate an approach to computing operational Availability for certain important special cases in which deployment period is much greater than mean time to repair μ and of the same order of magnitude as mean time between failures M.

COMPUTING AVAILABILITY

The first step in computing Availability is to define the term precisely and unambiguously in a manner which is consistent with accepted usage and the specific operational requirement.

While there may be good reasons for preferring a slightly different definition of the term, the need for general acceptance and the existence of a United States Government-promulgated[1] definition of the term would seem to encourage acceptance of the following definition:

Availability:

A measure of the degree to which an item is in the operable and committable state at the start of the mission, when the mission is called for at an unknown (random) point in time.

While this definition satisfies the need for general acceptance, it fails to satisfy our requirements for precision. For the purposes of making specific analyses, the author has found it convenient to interpret this definition in terms of a single demand which is equally likely to occur at any time during the deployment period.

Calabro[2] considered other special cases which are consistent with other interpretations of the term Availability. While the author is not aware of use of renewal theory to obtain Availability estimates as such, extension of the work of Bazovsky[3] and others to include maintenance factors would not seem to impose insurmountable mathematical difficulties. The incentive to compromise rigor by use of approximate explicit equations rather than to use the more general approach is convenience and cost.

While this paper is primarily concerned with situations in which there is at least some on-site repair capability, it is nevertheless convenient to start with the availability condition in which equipment repair is not possible and for which the survival probability is adequately represented by an exponential function.

$$p_o\{t\} = e^{-\frac{t}{M}} \tag{2}$$

Since Availability is defined as the probability that the equipment is operable at some random time (t) during the deployment period, such that

$$0 \leq t \leq T,$$

then the Availability is the average value of (2). That is,

$$A_o(T) = \frac{1}{T}\int_0^T e^{-\frac{t}{M}} \, dt \tag{3}$$

$$= \frac{M}{T}\left[1 - e^{-\frac{T}{M}}\right] \tag{4}$$

Before proceeding to cases in which on-site repair capability exists, we will consider the significance of (4) under conditions of interest. In particular, we consider the case in which an item having an MTBF of 20,000 hours is installed on a ship with a deployment period of 2,000 hours and no on-board maintenance capability exists. Substituting in (4),

$$A_o\,(2000) = \frac{20{,}000}{2000}\left(1 - e^{-\frac{2000}{20{,}000}}\right)$$

$$= 0.9516$$

For comparison, the probability that the unit will not fail during the deployment period is

$$R\,(2000) = e^{-\frac{2000}{20{,}000}}$$

$$= 0.9048$$

This value is not in conflict with the intuitive feeling that the probability of being able to meet a demand (which will occur, on the average, halfway through the deployment period) would approximate the reliability for half the deployment period (mission) duration.

Next, consider the situation which exists if n spare units (or enough parts to make n repairs) are carried. Under such circumstances, the criteria for success are:

- The unit does not fail during the period of deployment starting from time zero through the time of demand, or
- The unit fails as many as n times prior to the time of demand but was repaired each time.

The first step in the estimation of success probability is to estimate the probability having exactly i failures, hence the need for i spares in the first t hours of the deployment period. The actual operating time during this period is t hours, less the time required to make i repairs.

It is possible to make a mathematical model for any distribution of repair times; however, the quality of data rarely justifies a level of sophistication beyond that of a fixed repair time equal to μ, where both deployment period T and mean time between failure M are quite large compared to μ. In the cases of interest, μ has a range of a few minutes to a few hours, and neither M nor T are less than a few hundred hours.

Using the Poisson formula, the probability of having exactly i failures in the first t hours of the deployment period is

$$p\left\{\frac{t}{M}, i\right\} = \frac{\left(\frac{t - i\mu}{M}\right)^i \quad e^{-\left(\frac{t - i\mu}{M}\right)}}{i\,!} \tag{5}$$

Since we assume that each of the first n failures are repaired in exactly μ hours, the probability of mission failure is the probability that the logistics limit of n failures is not exceeded and that no repairable failure occurs in the μ hours just prior to the demand. The probability of failure in this time interval is

$$p\left\{t - \mu, t\right\} = e^{-\frac{\mu}{M}} \tag{6}$$

Combining (5) and (6) gives

$$p_n\left\{t\right\} = e^{-\frac{t}{M}} + e^{-\frac{\mu}{M}} \sum_{i=1}^{n} p\left\{\frac{t}{M}, i\right\} \tag{7}$$

In practical cases, the error introduced in (5) by use of t rather than $t - i\mu$ is too small to justify the additional labor required to obtain useful information from field data of usual quantity and accuracy[4].

Equation (7) is expanded to show the important special cases of n = 1, 2, 3 and ∞ as well as the case for n = 0 which was shown in (2).

$$p_o\left\{t\right\} = e^{-\frac{t}{M}} \tag{8}$$

$$p_1\left\{t\right\} = e^{-\frac{t}{M}}\left[1 + e^{-\frac{\mu}{M}}\left(\frac{t}{M}\right)\right] \tag{9}$$

$$p_2\left\{t\right\} = e^{-\frac{t}{M}}\left[1 + e^{-\frac{\mu}{M}}\left\{\left(\frac{t}{M}\right) + \frac{1}{2}\left(\frac{t}{M}\right)^2\right\}\right] \tag{10}$$

$$p_3\left\{t\right\} = e^{-\frac{t}{M}}\left[1 + e^{-\frac{\mu}{M}}\left\{\left(\frac{t}{M}\right) + \frac{1}{2}\left(\frac{t}{M}\right)^2 + \frac{1}{6}\left(\frac{t}{M}\right)^3\right\}\right] \tag{11}$$

$$p_\infty\left\{t\right\} = e^{-\frac{\mu}{M}} + e^{-\frac{t}{M}}\left(1 - e^{-\frac{\mu}{M}}\right) \tag{12}$$

The probability functions (8) through (12) may now be used to compute Availability values for a specific deployment period T, in which demand is in accordance with a uniform distribution by finding the average values of those functions.

For the sake of example, the required integration is carried out for n = 1 and for $n \to \infty$.

For n = 1

$$A_1\,(T) = \frac{1}{T}\int_o^T e^{-\frac{t}{M}}\left[1 + e^{-\frac{\mu}{M}}\left(\frac{t}{M}\right)\right] dt \tag{13}$$

$$= \frac{M}{T}\left(1 - e^{-\frac{T}{M}}\right) + e^{-\frac{\mu}{M}}\left[\frac{M}{T} - \left(\frac{M}{T} + 1\right) e^{-\frac{T}{M}}\right] \tag{14}$$

For $n \to \infty$

$$A_\infty\,(T) = e^{-\frac{\mu}{M}} + \frac{M}{T}\left(1 - e^{-\frac{\mu}{M}}\right)\left(1 - e^{-\frac{T}{M}}\right) \tag{15}$$

To appreciate that (15) is only approximation to the true Availability function, note that:

$$\lim_{T \to \infty} A_\infty\,(T) = e^{-\frac{\mu}{M}} \tag{16}$$

whereas the correct function for the case of infinited deployment and perfect support is (1).

For cases in which M is ordinarily several orders of magnitude greater than μ, (16) and (1) differ in the fifth significant figure so that this approximation is of no practical consequence. Since use of (1) avoids looking up values in a table, the author prefers to use the approximation (17) for most computations.

$$e^{-\frac{\mu}{M}} \doteq \frac{M}{M + \mu} \tag{17}$$

No special problems are encountered in computing equations for $A_2(T)$, $A_3(T)$, etc. However, there is another important special case for which a comment might be appropriate. That is the case in which a single spare is used as back up for (j) operation units. For this case, the equipment group mean time between failures $\frac{M}{j}$ is used in the Availability equation instead of the unit mean time between failures (M).

For the specific case in which one spare unit is used as back up for two operational units,

$$p_{1/2}\left\{t\right\} = e^{-\frac{2t}{M}}\left[1 + e^{-\frac{2\mu}{M}}\left(\frac{2t}{M}\right)\right] \tag{18}$$

$$A_{1/2}(T) = \frac{M}{2T}\left(1 - e^{-\frac{2T}{M}}\right) + e^{-\frac{2\mu}{M}}\left[\frac{M}{2T} - \left(\frac{M}{2T} + 1\right) e^{-\frac{2T}{M}}\right] \tag{19}$$

Conceptually similar methods may be used for other cases in which n spares are used to support j operational units; however, the equations tend to get out of hand for more complex situations so it may be advantageous to make use of the more general methods of renewal theory[2].

Another point to consider in the use of the Availability values is that, unlike the steady state case, it is not possible to compute the Availability of a series system by taking the product of the individual equipment Availabilities. Instead, it is necessary to take the product of the individual probability values such as are shown as (8) through (12), then to find the average value of the resulting function for the period from $t = 0$ to $t = T$.

To illustrate the magnitude of errors which can result from the use of the steady state availability equation for a typical shipboard situation, we choose some (perhaps) typical equipment and deployment period values:

Deployment Period (T) = 1000 hours
Mean time between failures (M) = 2000 hours
Mean time to Repair (μ) = 1 hour

Using the steady state equation (1),

$$A = \frac{2000}{2000+1}$$

$$= 0.9995$$

Using the more sophisticated Availability equation for the case of a single spare unit (14),

$$A_1(1000) = \frac{2000}{1000}\left(1 - e^{-\frac{1000}{2000}}\right) + e^{-\frac{1}{2000}} \left[\frac{2000}{1000} - \left(\frac{2000}{1000} + 1\right) e^{-\frac{1000}{2000}}\right]$$

$$= 0.9673$$

Using the equation for unlimited spares (15),

$$A_\infty(1000) = e^{-\frac{1}{2000}} + \frac{2000}{1000}\left(1 - e^{-\frac{1}{2000}}\right)\left(1 - e^{-\frac{1000}{2000}}\right)$$

$$= 0.9999$$

These results give an indication of the importance of using more sophisticated equations than (1) to compute operational Availability, especially for situations in which on-site spares are severely limited or equipment mean time between failures is greater than deployment period.

REFERENCES

1. Mil-Std-721, Definitions at Effectiveness Terms for Reliability, Maintainability, Human Factors, and Safety.
2. Calabro, S. R., RELIABILITY PRINCIPLES AND PRACTICES, McGraw Hill, 1962.
3. Bazovsky, I., "Reliability of Rifles and Machine Guns and Other Small Arms"
4. Jacks, H. G., "Total Confidence Limits on Observed Reliability," 1971 ANNALS OF THE ASSURANCE SCIENCES.

PERIL RATE/MTBF (t_1, t_2)

Harold Ascher

Naval Research Laboratory
Washington, D.C. 20390

1. Introduction

The purpose of this paper is to discuss the various mean times to (or between) failure(s) of interest in the reliability field. An understanding of why different models are needed for mean failure times of parts, sockets and systems depends on an understanding of the different reliability concepts which apply to these different hierarchical levels. The reader is referred to Ascher (1) and Ascher and Feingold (2) for a detailed treatment of these basic concepts. However, an outline of the most fundamental points made in these two references is presented below for ready reference. In addition, the notation and terminology to be used in the discussion of mean times will be established.

2. Basic Concepts

A. Definitions

The following definitions apply for the purposes of this paper.

(1) Part - An item which is not subject to disassembly and hence, is discarded when it fails.

Comment: It is not always clear-cut when it is feasible to disassemble an item. For example, while most vacuum tubes are discarded when they fail, some expensive microwave tubes are disassembled and restored to operating condition. For these microwave tubes, their subcomponents would be considered parts rather than the tube itself.

(2) Socket - A circuit or equipment position which, at any given time, holds a part of a given type; as parts fail, they are replaced by new or good-as-new parts from the same statistical population as the original part.

Comment: This definition is meant to include, but not be restricted to, actual physical sockets such as those which hold tubes or transistors. Other examples of what is meant by the generalized definition of a socket follow:

a. the position between two terminals on a breadboard, as opposed to the particular part, say a resistor, which is installed between the two terminals at a given time.

b. the position in an engine head which holds successive spark plugs. It is assumed that either a new plug is installed whenever the previous one fails or the failed plug is cleaned thoroughly so as to be indistinguishable from new. If this "good-as-new" requirement does not hold, then the head position is not a socket under the above definition. In the case of physical sockets, it is possible for the socket itself to fail. Since most sockets are designed to fail much less frequently than the parts they hold, the possibility of socket failure will be disregarded. The cumbersome alternative would be to consider the physical socket itself to be held in a generalized socket.

(3) System - A collection of two or more sockets and their associated parts interconnected in such a way as to perform one or more functions.

(4) Non-repairable system - A system which is discarded when it ceases to perform satisfactorily.

Comment: An example of a non-repairable system is an unmanned satellite which has no provision for remote switching of redundant circuits. It will be noted that though a non-repairable system contains at least two parts, it is indistinguishable from a part for some purposes. For example, if an unmanned satellite fails and the function it performs is to be continued, another satellite must be orbited. If the replacement satellite is from the same statistical population as the failed one then, in a sense, it is the replacement part in the "socket in the sky."

(5) Repairable system - A system which, after failure to perform at least one of its functions, can be restored to performing all of its required functions by the replacement of, at most, some of its constituent parts.

Comment 1: The above definition is worded to include the possibility that no parts are replaced. For example, the system might be repaired by an adjustment or by a well directed kick.

Comment 2: A system which has redundant paths which are repaired but which is discarded as soon as it fails to perform at least one of its required functions is not considered a repairable system.

B. Basic mathematical models for parts, sockets and systems

(1) Distribution function of time to failure of a part.

When T = random variable, time to failure

$F_T(t) \equiv \Pr\{T \le t\} \equiv$ distribution function of T

$f_T(t) \equiv \frac{d}{dt} F_T(t) \equiv$ density function of T, when this derivative exists for all $t \ge 0$

$R_T(t) \equiv 1 - F_T(t) \equiv \Pr\{T > t\} \equiv$ reliability function of T

$$h_T(t) \equiv \frac{f_T(t)}{R_T(t)} \equiv \text{hazard function of } T$$

$$h_T(t)\,dt = \Pr\{t \leq T \leq t + dt \mid T \geq t\} \qquad (1)$$

Another way of stating that a group of parts are from the same statistical population is to indicate that when they are operated under identical conditions and have identical failure criteria, they have identical distributions of time to failure, $F_T(t)$, or equivalently, that they have identical reliability functions, $R_T(t)$. Then equation (1) can be interpreted as follows: $h_T(t)\,dt$ is the conditional probability that a part (from the population with distribution of time to failure $F_T(t)$) put into service at $t = 0$, and known to have operated until t, fails in $(t, t + dt)$.

(2) Renewal Process as a model of times between successive socket failures.

A renewal process is defined as a sequence of independent non-negative identically distributed random variables not all 0 with probability 1. If we let T_1, T_2, T_3, ..., T_i, ... be this sequence of random variables, which in our context, will be times-between-successive-failures (interarrival times), and we let

$F(t) \equiv$ the distribution of each of the T_i's

$f(t) \equiv F'(t) \equiv$ the probability density function (pdf) of each of the T_i's

$F^{(K)}(t) \equiv$ the distribution of $\sum_{i=1}^{K} T_i$

$f^{(K)}(t) =$ the p d f of $\sum_{i=1}^{K} T_i$

and we define $F^{(0)}(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}$

then it can be shown (Barlow and Proschan (3)) that

$$\Pr\{N(t) = n\} = F^{(n)}(t) - F^{(n+1)}(t)$$

where $N(t)$ is the number of renewals (in our context, the number of failures) in the interval $(0, t)$.[1] If we now define $M(t)$ to be the expected number of renewals in $(0, t)$

$$M(t) \equiv E\{N(t)\}$$

then Barlow and Proschan show that

$$M(t) = \sum_{K=1}^{\infty} F^{(K)}(t)$$

[1]It is assumed here and throughout the paper that repair times are either instantaneous or measured on a different time scale.

The derivative of $M(t) \equiv r(t) \equiv M'(t)$

and $r(t) \equiv \sum_{K=1}^{\infty} f^{(K)}(t)$

where $r(t)$ is the renewal rate.

Since the previously introduced definition of a socket requires that the successive installed parts are nominally identical (in the sense that they will have identically distributed failure times when equally stressed), a renewal process is a very plausible model for the sequence of failure times in the socket[2]. In this context, $r(t)\,dt$ can be interpreted as the (unconditional) probability that the part in the socket at time t, fails in the interval $(t, t + dt)$. Since the part in the socket at t may be the first, second, third, etc., part installed in the socket from the time the first one was installed at $t = 0$, the failure in $(t, t + dt)$ will correspondingly be the first, second, third, etc., failure in the socket. In the special case where the interarrival times are exponentially distributed, it can be shown that the renewal rate becomes a constant, λ, which is numerically equal to the hazard function of each of parts installed in the socket. Even in this special case, however, there is no equivalence between renewal rate and hazard function since the condition which results in numerical equality does not alter the fundamental differences in the way these two terms are defined. That is, since the hazard function is the intensity with which one part is tending to fail, while the renewal rate is the time derivative of an expected number of failures, they can never become equivalent. The prime reason that they are often erroneously thought to be equivalent is that each has often been called failure rate. The misleading effect of calling the hazard function "failure rate" may perhaps be better understood by turning to a different context. In the maintainability field the hazard function is equally misleadingly called "repair rate" and is expressed as, say, 2 repairs per hour. This conjures up an image of a repairman turning out repairs at an average rate of 2 per hour - even when he has only one repair to perform. The necessary distinction between single and multiple events is better understood in the field of Queuing Theory where the number of arrivals per unit time is modeled by the rate associated with a stochastic process and the intensity with which each service time is tending to be completed is modeled by the hazard function of the appropriate service time distribution.

(3) Poisson Process as a model of the number of system failures in a given time. In order to define a Poisson Process we

[2]It should be noted, however, that a renewal process is not automatically the correct model for a socket. For example, the stress on the part in the socket may change over the course of time, hence changing the distribution of time to failure even for nominally identical parts.

must first introduce the notion of a counting process. A stochastic process $\{N(t), t \geq 0\}$ is said to be a counting process if $N(t)$ represents the total number of events which have occurred in the interval $(0, t)$. The counting process $\{N(t), t \geq 0\}$ is said to be a homogeneous Poisson Process if

(i) $N(0) = 0$

(ii) $\{N(t), t \geq 0\}$ has independent increments

(iii) The number of events (in our context, failures)

in any interval of length $t_2 - t_1$ has a Poisson distribution with mean $\rho(t_2 - t_1)$. That is, for all $t_2 > t_1 \geq 0$,

$$\Pr\{N(t_2)-N(t_1)=n\} = \frac{e^{-\rho(t_2-t_1)}\{\rho(t_2-t_1)\}^n}{n!} \tag{2}$$

for $n \geq 0$.
From (2) it follows that

$$E\{N(t_2 - t_1)\} = \rho(t_2 - t_1)$$

where the constant, ρ, is the rate of occurrence of failures. The present author has called the Poisson process' rate of occurrence the "peril rate" in previous papers and this nomenclature will be retained here.

It can be shown that the successive times between failures of the homogeneous Poisson process defined above are independent and identically distributed exponential random variables. Hence, the homogeneous Poisson process is a special case of a renewal process. Therefore, it is an appropriate model for a socket containing parts with exponentially distributed times between failures. It is also the correct model for the times between failures of a series repairable system each of whose sockets contains exponential parts or, under mild restrictions, for the times between failures of an "infinitely" complex system which has been operating "infinitely" long (Drenick (4)). When these conditions are not met, at least approximately, the homogeneous Poisson process will not be an appropriate model. In the case of repairable systems other renewal models will usually not be appropriate either, since most repairs involve the replacement of only a small fraction of the system's parts. Hence, even if these repairs restore performance to original specifications they do not renew the system in the reliability sense, since after repair most parts retain their full age. If we go to the other extreme and assume that each repair renews the system in the performance sense but leaves it with its full age ("bad-as-old") in the reliability sense, then it is shown in Ascher and Feingold (2) that a nonhomogeneous Poisson process is the appropriate model. (This reference and Ascher (1) should be consulted for an elaboration of the above points.)

The nonhomogeneous Poisson process alluded to above differs from the homogeneous one only in that the peril rate is a function of time rather than a constant. That is, conditions (i) and (ii) are retained and condition (iii) is modified to be

(iii) The number of failures in any interval (t_1, t_2) has a

Poisson distribution with mean $\int_{t_1}^{t_2} \rho(t)\,dt$.

That is, for all

$t_2 > t_1 \geq 0$,

$$\Pr\{N(t_2)-N(t_1)=n\} = \frac{e^{-\int_{t_1}^{t_2}\rho(t)dt}\{\int_{t_1}^{t_2}\rho(t)dt\}^n}{n!} \tag{3}$$

for $n \geq 0$.
From (3) it follows that

$$E\{N(t_2)-N(t_1)\} = \int_{t_1}^{t_2} \rho(t)dt$$

or for $t_1 = 0$

$$E\{N(t_2)\} = \int_0^{t_2} \rho(t)dt$$

since $N(0) = 0$.

The interpretation of $\rho(t)\,dt$ is that it is the probability that a system put into service at $t = 0$ and repaired in a "bad-as-old" sense fails in the interval $(t, t + dt)$. It is hardly surprising that this is a similar interpretation to $r(t)\,dt$ since in the special case of a homogeneous Poisson process $\rho(t) \equiv r(t) = \lambda$. It is stressed that just as for the renewal rate, the relationship between a constant peril rate and the hazard function of the corresponding exponential distribution of interarrival times is one of numerical equality rather than of equivalence.

It is emphasized that even a nonhomogeneous Poisson Process is not necessarily an appropriate model for a repairable system. For example, a system composed of redundant paths which are not repaired until system failure occurs, will not satisfy the requirement of independent increments (condition ii). That is, the number of this system's failures in one interval will not, in general, be independent of the number of system failures in a preceding, non-overlapping interval (if no system failures occur in one interval, the probability of at least one system failure in the next interval is increased). The complex stochastic processes needed to model such systems will not be considered in this paper.

3. Models for the Mean Time to Failure (MTTF) of a part

A. Probabilistic models

A given population of parts, operated under equal stresses, will have a distribution of time to failure $F_T(t)$. Then the MTTF is simply the mean of that distribution:

$$\text{MTTF} \equiv \int_0^{\infty} t \, d\, F(t) \tag{4}$$

whenever this improper integral exists. Equivalently,

$$\text{MTTF} \equiv \int_0^{\infty} t \, f(t) \, dt$$

when the distribution has a probability density function. Evans (5) shows that whenever the integral in equation (4) exists then

$$\text{MTTF} = \int_0^{\infty} R(t) \, dt$$

where $R(t) \equiv 1-F(t) \equiv$ the reliability function. In the special case where

$f(t) = \lambda e^{-\lambda t}$, $\text{MTTF} = 1/\lambda$, a well known result for the exponential distribution. Since for this distribution the hazard function, $h(t) = \lambda$, in this special case the MTTF and hazard function are reciprocals. It is often thought that a similar relationship always holds, i.e., that the MTTF always equals the reciprocal of the average hazard function $\overline{h(t)}$ where

$$\overline{h(t)} = \lim_{t \to \infty} \frac{1}{t} \int_0^t h(x) \, dx$$

That this is not the case is demonstrated by a simple counterexample. For the two parameter Weibull distribution,

$$F(t) = 1 - e^{-\lambda t^{\alpha}} \qquad \lambda, \alpha > 0$$

$$h(t) = \lambda \alpha t^{\alpha - 1}$$

$$\overline{h(t)} = \lim_{t \to \infty} \frac{1}{t} \int_0^t \lambda \alpha x^{\alpha - 1} \, dx$$

$$= \lim_{t \to \infty} \lambda t^{\alpha - 1} = \begin{cases} 0, & \alpha < 1 \\ \lambda, & \alpha = 1 \\ \infty, & \alpha > 1 \end{cases}$$

Since the expected value of a Weibull distributed random variable is

$$E\{T\} = \frac{1}{\lambda} \Gamma\left(1 + \frac{1}{\alpha}\right)$$

it is apparent that $\overline{h(t)} = \dfrac{1}{E\{T\}}$ only in the special case where $\alpha = 1$.

B. Statistical Models

The problem we are considering here is estimating the mean of a known, or unknown, distribution. In the case of many known distributions, optimum estimators are available. For example, if times to failure are exponentially distributed and n times to failure $T_1, T_2, \ldots T_n$ are observed then a minimum variance unbiased estimator of MTTF is

$$\widehat{\text{MTTF}} = \frac{\sum_{i=1}^{n} T_i}{n},$$

the sample mean. The same result holds for the untruncated normal distribution and if the truncation is slight, it will be close to optimum. In the (practically speaking, universal) case where the distribution of time to failure is not exactly known, the sample mean will not necessarily be optimum but it usually will provide results which would not be greatly improved even if the distribution were known. It is also reassuring to recall that for a distribution with finite mean the strong law of large numbers states, that with probability one, the sample mean will converge to the true mean as sample size increases. In the case of distributions which do not have finite mean , the sample mean may be of no more value as an estimator of central tendency than a single sample but these distributions are not applicable to physical quantities such as time to failure.

4. Models for the Mean Time Between Failures (MTBF) of a socket

A. Probabilistic models

By assumption, we are seeking the mean of the interarrival times of a renewal process. Intuitively we would expect that the MTBF of the socket would equal the MTTF of each of the parts installed in the socket. In practice, however, the socket MTBF depends on whether we begin observing the socket at the time that a new part is installed or at some arbitrary time when the age of the part in the socket is unknown. (Conceivably, we could start our observation of the socket at a known time after installation of the part presently in the socket. Since this is an artificial situation that would complicate matters, it will be ignored). When we start observing at an arbitrary time, Cox and Lewis (6) show that,

$$\text{MTBF} \equiv \frac{t_o}{M(t_o)} = \mu,$$

where t_o is the period of observation and μ = MTTF of each of the parts in the socket. However, when we start observing at the time that a new part is inserted in the socket, this result is only asymptotically true, i.e.,

$$\text{MTBF}_{\infty} \equiv \lim_{t_o \to \infty} \frac{t_o}{M(t_o)} = \mu$$

where the subscript is used to denote the asymptotic or steady state value. The reason that this is so can be seen from the following example: assume that the new part, installed at the time observation begins, has a distribution of time to failure with a large MTTF and a small variance. Assume further that the total time observation, t_o, is small compared to the MTTF. Then the probability that even one failure occurs in $(0, t_o)$ is very small and the ratio $\frac{t_o}{M(t_o)}$ will tend to overestimate the MTTF of the part, and hence, the MTBF of the socket.

B. Statistical models

It has been indicated earlier that a renewal process is a very plausible model for a socket. Therefore, in a probabilistic model the assumption of renewal is the only reasonable one to make. Nevertheless, when data are available the renewal hypothesis can be and should be checked. This is done by testing the successive interarrival times from one socket for any tendency for these times to increase (decrease). Two possible reasons for these times to increase (decrease) are improvement (reduction) in the quality of spare parts and milder (more severe) stress on the socket with the passage of time. Statistical tests to detect such trends are described by Cox and Lewis (6) and Mann (7), and applied to reliability problems by Bassin (8) and Ascher and Feingold (2). Assuming that no statistically significant trend is demonstrated, the socket MTBF can be estimated as follows. Cox and Lewis (6) show that if we start observing from the time a new part is installed in the socket and observe until the n^{th} failure has occurred then the sample mean, $\frac{1}{n}\sum_{i=1}^{n} T_i$, is an unbiased estimator of the MTBF. If we begin observations at an arbitrary time, the sample mean is biased by the amount

$$\frac{C^2(T)\,\mu}{n} \sum_{k=1}^{n} c_k$$

where C is the coefficient of variation of $F_T(t)$ and the c_k are defined as

$$c_k \equiv \frac{\text{cov}(T_i, T_{i+k})}{\text{var}(T)} \quad (k = \ldots -1, 0, 1 \ldots)$$

In general, this bias tends to zero as the number of samples increase.

5. Models for the MTBF (t_1, t_2) of a system

A. Probabilistic models

From the viewpoint of their mathematical treatment systems can be separated into two categories, repairable and nonrepairable. Obtaining the distribution of time to failure for a nonrepairable system as a function of 1) the distributions of its constituent parts and 2) the manner in which the parts are interconnected, may be a complicated procedure but once this distribution is obtained the MTTF can be calculated by equation (4) just as for a part. If another nonrepairable system from the same population is installed to continue performing the function of the first, then the two systems, together with succeeding ones, are analogous to parts in a socket and the MTBF of the "socket" can be calculated by the methods of the previous section. When we consider repairable systems, the situation becomes much more complex since such a system contains parts of mixed ages once replacement parts are installed. Barlow and Proschan (9) have recently presented an asymptotic result for a series system. They have shown that

$$\text{MTBF}_{\infty} = \frac{1}{\sum_{i=1}^{K} \frac{1}{\mu_i}}$$

where $0 < \mu_i < \infty$ is the MTTF of the i^{th} component of a system containing K parts[3]. The subscript ∞ denotes that this is a steady state result. The above result holds regardless of the distributions of the times to failure of the parts comprising the system (except for the condition that the mean must be finite). Of course, in the special case where each part has exponentially distributed time to failure, the above result holds from the initial time of system operation rather than just in the steady state.

The existence of this result does not close out the problem of quantifying the average time between failures of even a series system. The most obvious problem is the asymptotic nature of the result; for example, a system may be discarded because of technological obsolesence before it has operated long enough for this result to be an adequate approximation. As discussed earlier in this paper, a nonhomogeneous Poisson Process may be a reasonable model for systems which have not achieved the steady state. This model is flexible enough to handle both repairable system burn-in and wearout.

A decreasing peril rate models a reliability growth situation where the number of failures per unit time is decreasing because of burn-in, design fixes,

[3] This result holds almost surely, i.e., other results are possible but their probability of occurrence is zero.

learning curves for operators or maintenance men, etc. There have been many papers published which have unconsciously used this model while using terminology like "failure rate", "cumulative hazard rate" and "MTBF." The last term, which implies a single fixed value, is used even though the emphasis is on how consecutive times between failures are tending to increase. It is proposed that the term be modified to MTBF (t_1, t_2), that is, the mean time between failures over the interval (t_1, t_2) where $t_2 > t_1 \geq 0$. This notation will be used in this paper.

A nonhomogeneous Poisson Process with an increasing peril rate is an appropriate model when a repairable system is wearing out. For example, it is a widely accepted model for an automobile since this system's age is intuitively measured from the time it was first put into service. It is also noted that popular conceptions about automobiles are probably the cause of the widespread acceptance of the bathtub curve as a model for the rate of occurrence of failures of repairable systems. This is in spite of Drenick's Theorem (4) which states that this rate should asymptotically approach a constant. This inconsistency results from the fact that the rapid rise in the peril rate at about 100,000 miles causes the car to be scrapped, thus preventing the prolonged period of operation required for Drenick's Theorem to apply.

In the case of either decreasing or increasing peril rates, the expected number of failures $E\{N(t_1, t_2)\}$ of a system in (t_1, t_2) given that the system began operation at $t = 0$ is:

$$E\{N(t_1, t_2)\} = \int_{t_1}^{t_2} \rho(t)\, dt$$

Then:

$$\text{MTBF}(t_1, t_2) \equiv \frac{t_2 - t_1}{\int_{t_1}^{t_2} \rho(t)\, dt}$$

The notation MTBF (t_1, t_2) also would be appropriate for a system modeled by any other stochastic process with independent increments, i.e., any other Markov Process. However, for systems which cannot be modeled by a process with independent increments, the MTBF in an interval (t_1, t_2) is not independent of what occurred in $(0, t_1)$ and in general, it would be very difficult to obtain an expression for

$$\text{MTBF}\{(t_1, t_2) \mid \text{history over } (0, t_1)\}.$$

For such a system, e.g. a repairable, redundant system whose redundant elements are not repaired as they fail, an alternative approach must be adopted. The Mean Time to First System Failure, MTFSF, of such a system can be calculated by Equation (4) just as for the MTTF of a part. The MTFSF does not give as much information about the repairable system as MTTF gives for a part, but the MTFSF is still a useful parameter.

B. Statistical Models:

In the section on the statistical treatment of sockets, it was stated that when data are available the hypothesis of renewal should be tested, rather than assumed a priori. The reverse comment applies to repairable systems. When data are available for repairable system interfailure times, a trend should not be assumed a priori. Rather, a null hypothesis of a renewal process, or perhaps more specifically, a homogeneous Poisson Process, should be tested against the alternative of trend. If no trend is established, the methods of Section 4B should be applied. If a trend exists, other methods must be used. One procedure is to assume a model for the peril rate and then test it for goodness of fit as shown in Ascher and Feingold (2). If the model for peril rate, $\rho^*(t)$, is accepted, then:

$$\widehat{\text{MTBF}}(t_1, t_2) = \frac{t_2 - t_1}{\int_{t_1}^{t_2} \rho^*(t)\, dt}$$

A more provisional method is to count the number of failures in $(t_1, t_2) = \hat{N}(t_1, t_2)$.

Then the MTBF (t_1, t_2) is estimated by:

$$\widehat{\text{MTBF}}(t_1, t_2) = \frac{t_2 - t_1}{\hat{N}(t_1, t_2)}$$

The advantage of this estimator is that it is easily calculated. Obviously, it is subject to large sampling fluctuations, particularly for small values of $t_2 - t_1$.

It is noted that many proponents of reliability growth use plots of $\widehat{\text{MTBF}}(o, t)$ versus time (t) or estimated peril rate versus time to "demonstrate" that reliability growth is actually taking place. This method is highly subjective and should be replaced by formal tests for trend. It has been suggested by Sessen (10) that "the" cause of the apparent increase of MTBF (o, t) over long periods of time is due to the influence of the relative large number of failures which occurred for small values of t. That is, if reliability growth does take place over a period of time, say (o, t^*), then even if the peril rate is a constant after t^* the effect of the early failures does not become negligible until $t \gg t^*$. The possibility that growth does not continue after t^* can be checked by conducting a test for trend on only failure times which occur after t^* where t^* might

be selected on the basis of the history of implementation of design fixes or past experience. One drawback of this approach is that trend tests have low power (i.e., poor ability to reject the null hypothesis of no trend when it is not true) for the sample sizes usually encountered. When some of the data are censored the power will, in general, be reduced further. Hence, acceptance of trend for all the data and rejection of trend for the data after t^* may be an indication of low power rather than absence of trend after t^*. While this is a real limitation of this procedure, this limitation is by no means unique to this situation. As more data are accumulated, the trend test can be repeated to recheck the hypothesis that reliability growth is continuing indefinitely.

References

1. H. Ascher, "Hazard Functions, Renewal Rates and Peril Rates," Proceedings of the Ninth R&M Conference, 1970, Society of Automotive Engineers, New York.

2. H. Ascher and H. Feingold, "'Bad-As-Old' Analysis of System Failure Data," Proceedings of the Eighth R&M Conference, 1969, Gordon and Breach, New York.

3. R. Barlow and F. Proschan, "Mathematical Theory of Reliability," John Wiley, 1965.

4. R. Drenick, "The Failure Law of Complex Equipment," J. SIAM, Volume 8, Number 4, Pages 680-690.

5. R. Evans, "Proper Proof of a Reliability Theorem," IEEE Transactions on Reliability, Vol. R-18, Number 4, pp 205-206, Nov. 1969.

6. D. Cox and P. Lewis, "The Statistical Analysis of Series of Events," John Wiley, 1966.

7. H. Mann, "Nonparametric Tests Against Trend," Econometrica, Volume 13, pp 245-259, 1945.

8. N. Bassin, "Increasing Hazard Function and Overhaul Policy," Proceedings of the 1969 Annual Symposium on Reliability, IEEE Catalog No. 69 C 8-R.

9. R. Barlow and F. Proschan, "Availability Theory for Multicomponent Systems," ORC 72-8, University of California, April 1972.

10. L. Sessen, "'Weibull' Decreasing Hazard Rate - Help or Hoax?," Proceedings of the 1971 Annual Symposium on Reliability, IEEE Catalog No. 71 C 2-R.

RELIABILITY DATA PLOTTING USING THE PEARSON CURVES

INDEX SERIAL NUMBER — 1077

By
Francis J. Brewerton
Louisiana Tech University

R. Wayne Gober
Louisiana Tech University

Philip F. Rice
Louisiana Tech University

One of the basic purposes of failure data collection and analysis is the eventual prediction of failure patterns. Since reliability is defined as the probability that a device will operate satisfactorily for a specified period of time, collection and analysis of failure data play an important role in determining this probability of successful operation. All reliability computations are based either directly on indirectly on an assumed or observed distribution of failures. However, conventional methods of data treatment and description are not always efficient or sufficiently accurate for determination of failure probabilities.

Conventional Data Description

Failure data may be treated as having either a known probability distribution or an unknown probability distribution. Although failure data may be treated in the former manner, very few components' failure distributions are truly known. The more common practice is to find a theoretical distribution which reasonably approximates the actual observed data and then to treat the observed data as having come from that distribution. Commonly used theoretical distributions in reliability work include but are not limited to the normal, beta, gamma, Weibul, exponential, and log-normal distributions. Once the theoretical distribution has been assumed, some determination of whether or not the observed data contridict the assumed model must be made.

Probability Plotting

Probability plotting is a subjective method for testing the assumed theoretical distribution's descriptive ability in that the determination regarding the distribution's appropriateness is based on a subjective visual examination. The typical procedure is to plot data on special graph paper designed for the particular assumed theoretical distribution. If the model is adequate, the plot of the data will be approximately linear and percentiles and parameter values may then be estimated. If the plot is not sufficiently linear to satisfy the analyst, then a trial and error approach may be taken by plotting the same data on probability paper for other well-known distributions until a plot is obtained which is sufficiently linear. However, the determination of what can or cannot be considered a linear plot remains a subjective matter and two people analyzing the same plot might easily arrive at different conclusions.

Goodness of Fit Tests

Statistical tests of distributional assumptions provide a more objective technique for determining the adequacy of data description. The conventional approach is to

1. Array and classify the data.
2. Select a theoretical distribution form.
3. Select and calculate a test statistic from the observed data.
4. Determine the probability of obtaining the calculated test statistic given the selected distribution.
5. Accept or reject the distribution as an adequate descriptor based on a comparison of the computed statistic and the table statistic at some level of significance.

A variety of statistical tests is available to evaluate distributional assumptions. Some of the more popular ones include the Chi-Squared tests, the W tests, and the Kolmogorov-Smirnoff test. The details of these tests are included in most statistics texts. Of these tests, the Chi-Squared is the oldest, most versatile, and most commonly used test since it is applicable to any distributional assumption. It's major disadvantage is that it is not a particularly powerful test, a feature resulting from its lack of sensitivity in detecting inadequate descriptions when relatively few observations are available.

Generalized Methods of Curve Fitting

A variety of general techniques for representing data is available. The most common of these generalized techniques include the Johnson and Pearson distributions, the Gram-Charlier series, the Edgeworth series, and curve fitting using the least squares and maximum likelihood techniques.

The method of least squares involves the adjustment of observations so that the sum of the squares of the differences between the actual facts and the adjusted figures is a minimum. Although used extensively in regression analysis, the method of least squares is sometimes weak in that it can lead to equations which are incapable of solution. However for many applications the method of least squares can produce results equally as good as any other generalized technique of curve fitting.

The "maximum likelihood" technique was developed by Fisher who used it largely to approximate a particular class of curves. The general difficulty in applying the method is its lack of soluble equations except for specialized cases. When unsolvable equations are encountered the equation constants must be determined by approximation.

The Gram-Charlier Type A series, Charlier's later B and C series, and Edgeworth's series are all generalized curve fitting techniques of some similarity. The strongest objection to using these techniques for describing actual data is that these series techniques may give negative frequencies, particularly near the tails; further, the series may behave in an irregular sense, the sum of (k - 1) terms sometimes providing a better fit than the sum of k terms. None of these techniques has achieved any degree of popularity. Furthermore, none of the techniques mentioned here provides the degree of flexibility needed to describe the complete variety of forms assumed by the frequency distributions encountered in actual experience. However, both the Pearson family of curves and the Johnson distributions provide reasonable representations of observational data as well as approximations of theoretical distributions from known moments.

The Johnson Distributions

The Johnson distributions are empirical distributions based on transformations of a standard normal variate. One practical advantage of generating dis-

tributions in this fashion is that estimates of the percentiles of the fitted distribution can be obtained using a table of areas under a standard normal distribution. The Johnson distributions are categorized into three families designated as the S_L, S_U, and S_B forms. The S_L form is a three-parameter form while the S_U and S_B forms are four-parameter types. The general form of the transformation is

$$z = \gamma + \eta \, \tau (x; \varepsilon, \lambda) \quad \eta, \lambda > o; - \infty < \gamma, \varepsilon < + \infty \quad (1)$$

in which τ is an arbitrary function, γ, η, ε, and λ are four parameters of choice and z is a standard normal variate. The three forms S_L, S_U, and S_B are all obtained from equation (1). A Johnson distribution can be fitted with relative ease using equation (1) by determining which of the three distributional forms is applicable, estimating the parameters of the chosen family, and then obtaining the expected frequencies for the fitted distributions. The Johnson distributions include all curve shapes and provide descriptions equally as good as Pearsonian curves.

The Pearson Distributions

Karl Pearson devised a family of distributional curves, all of which emanated from the solution of the differential equation,

$$\frac{dy}{dx} = \frac{y(x + a)}{b_o + b_1 x + b_2 x^2} \qquad (2)$$

for the random variable x with probability density function y. The equation involves the four parameters a, b_o, b_1, and b_2.

The Pearson system includes twelve types of curves in addition to the normal curve. Types I, IV, and VI are the main types, with all other types being either transitional or trivial forms. All of the Pearson curves are fully determined by the first four moments, with some of the degenerate types being determined by fewer moments. Pearson's method of fitting the curves to observed data consists of

1. Determining the values of the first four moments of the observed distribution.
2. Calculating the observed values of β_1, β_2 and the criterion k (Pearson criterion) of determining the type to which the observed distribution belongs.
3. Equating the observed moments to the moments of this type of distribution expressed in terms of its parameters.
4. Solving the resulting equations for those parameters, whereupon the fitted distribution is determined.

In general, the Pearson curves do adequately overcome the practical difficulties associated with roughness of data, number of constants (and hence number of moments), and lack of systematic approach to data description. However, the Pearson system of curves requires a degree of mathematical sophistication and facility that is beyond many practitioners. Consequently, the simpler, less rigorous methods of data description have been typically used. This primary objection to the use of Pearson curves (or any other generalized system) can be largely overcome by computerizing the process thereby eliminating the mathematical tedium and sophistication. A computer program has been developed by Howe and Van Horn which makes Pearson curve fitting a practical and efficient method for describing data with unknown distributional form.

Program Features

The Howe-Van Horn program is a highly efficient method for numerically fitting any of the Pearson curves, creating random variables from the distribution estimating probabilities from the distribution, and evaluating the density at any point in its domain. The program also plots the density function over an interval of plus and minus six standard deviations.

Data may be introduced into the system by card or in blocks, with a maximum number of 3000 data points. The program will also fit a Pearson curve if the first four moments are supplied to the system, since these four moments completely determine the coefficients in equation (2). The moments may be taken about the mean or zero and any form of data or moment input may be used.

Curve Types

The family of solutions to equation (2) is generally separated into three main types of curves, depending on the nature of the roots of the quadratic in the denominator of the differential equation. According to Pearson, if the roots are real and of opposite sign, a Type I curve is indicated. The Type I curve includes some of the beta distributions. If the roots are real and of the same sign, a Type VI curve is indicated. The Type VI curve includes Snedecor's F distributions. If the roots are complex conjugate roots, a Type IV curve is indicated. The Type IV curve includes Student's t-distribution.

In addition to these main Types (I, IV, and VI), the Howe-Van Horn program includes the Pearson Type III curve and the normal curve. If the quadratic in the denominator of the differential equation reduces to a linear term, a Type III curve is indicated. The Type III curve includes the Chi-Square, gamma, and exponential distributions. If the quadratic reduces to a constant, a normal curve is indicated. All other Pearson Types (II, V, VII, VIII, IX, X, XI, XII) are no more than special cases of the Types I, III, IV, VI and the normal curve (Pearson's Type 0) and are handled automatically when they occur.

Although Pearson took some effort to classify his family of curves into different types, from the practical viewpoint of data description the classification of descriptive curves into Pearsonian types is somewhat academic. The appropriate Pearson type for a given set of data generally will be of secondary interest and importance in reliability work.

Description of Data from Unknown Distributions

The power of the Howe-Van Horn Pearson data description can be best illustrated through examples. Suppose the data shown in Table 1 represent grouped failure data for a given component for which no previous failure data exist. In order to compute reliabilities, some insight must be gained into the nature of the distribution of failures. Possible approaches to determining the distribution of the data include those discussed above. A typical approach might be to select class intervals and construct frequency distributions for these intervals. The optimum number and size of the class interval exist only in the mind of the observer, but two realistic possibilities are reproduced in Table 1 and Figure 1 and Table 2 and Figure 2. Either of the distributional representations shown in Figures 1 and 2 might be considered a reasonable description of the data. Goodness of fit tests could then be undertaken for one (or several) theoretical distributions to determine how closely the observed data conform to the theoretical distribution. Hopefully, one of the theoretical distributions will describe the observed data with sufficient conformity to satisfy the analyst and reliability computations may then commence. An alternative procedure is to plot the data in Table 1 successively on different types of probability paper until a plot with sufficient linearity is obtained. In either procedure, the approach is essentially by trial and error and sufficiency depends upon subjective evaluations.

The uncertainty in these procedures can be elim-

inated if Pearson distributions are used to describe the data. If the data in Table 1 are supplied directly to the Howe-Van Horn program, a precise curve is fitted immediately.

TABLE 1

Feasible Grouping for Unknown Distribution

	CLASS		FREQUENCY
	UNDER	0.0	0.
0.0	UP TO	2.00	340.
2.00	UP TO	4.00	219.
4.00	UP TO	6.00	150.
6.00	UP TO	8.00	96.
8.00	UP TO	10.00	80.
10.00	UP TO	12.00	46.
12.00	UP TO	14.00	29.
14.00	UP TO	16.00	9.
16.00	UP TO	18.00	9.
18.00	UP TO	20.00	6.
20.00	UP TO	22.00	5.
22.00	UP TO	24.00	2.
24.00	UP TO"	26.00	1.
26.00	UP TO	28.00	5.
28.00	UP TO	30.00	1.
30.00	UP TO	32.00	0.
32.00	UP TO	34.00	2.
34.00	UP TO	36.00	0.
36.00	AND OVER		0.

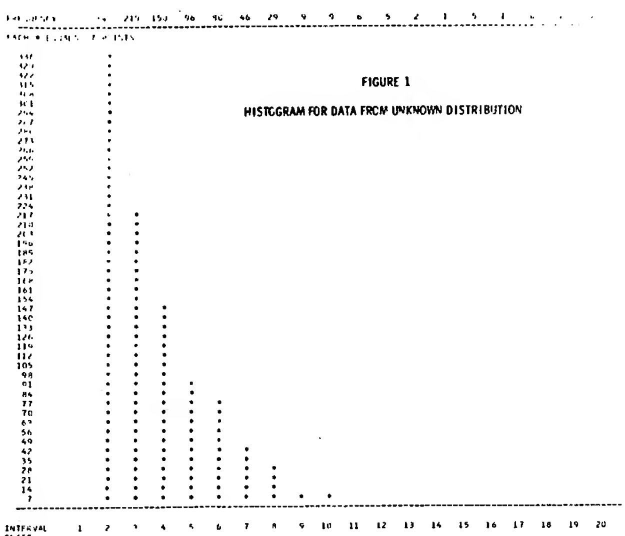


FIGURE 1

HISTOGRAM FOR DATA FROM UNKNOWN DISTRIBUTION

TABLE 2

Feasible Grouping for Unknown Distribution

	CLASS		FREQUENCY
	UNDER	0.0	0.
0.0	UP TO	4.00	559.
4.00	UP TO	8.00	246.
8.00	UP TO	12.00	126.
12.00	UP TO	16.00	38.
16.00	UP TO	20.00	15.
20.00	UP TO	24.00	7.
24.00	UP TO	28.00	6.
28.00	UP TO	32.00	1.
32.00	UP TO	36.00	2.
36.00	AND OVER		0.

Since the curve is determined by the first four moments of the data there is never any question regarding the accuracy of the data description. Figure 3 illustrates the fitted curve using Pearson data description for the data in Table 1. According to the program information supplied the Pearsonian equation for this curve is

$$f(x) = \text{Exp}\left[\log(.0773) - .1014\log\left(1 + \frac{x}{4.502}\right) + 65.75\log\left(1 - \frac{x}{334.5}\right)\right] \tag{3}$$

In actuality the data shown in Table 1 and Table 2 were generated from an exponential data generator to test the veracity of the Pearson curves for data description.

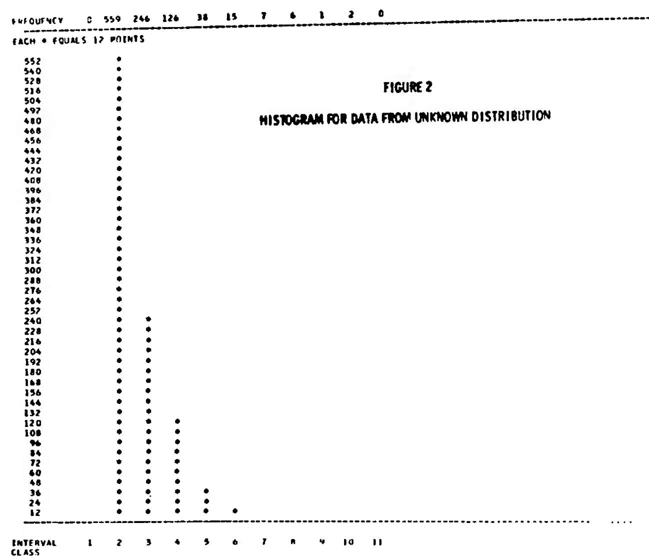


FIGURE 2

HISTOGRAM FOR DATA FROM UNKNOWN DISTRIBUTION

FIGURE 3 PEARSONIAN DESCRIPTION OF DATA FROM UNKNOWN DISTRIBUTION

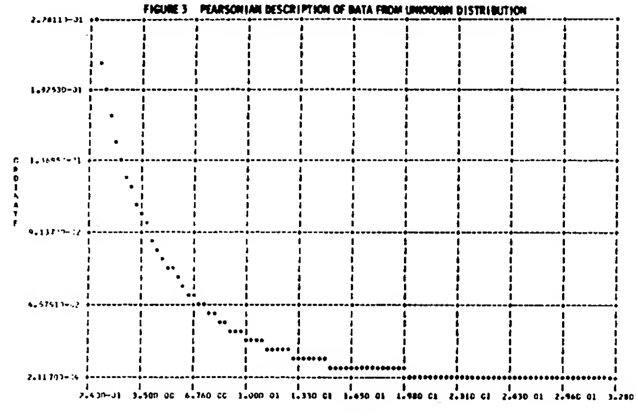


TABLE 3

Feasible Grouping for Unknown Distribution

	CLASS		FREQUENCY
	UNDER	0.0	0.
0.0	UP TO	0.25	85.
0.25	UP TO	0.50	152.
0.50	UP TO	0.75	184.
0.75	UP TO	1.00	194.
1.00	UP TO	1.25	170.
1.25	UP TO	1.50	90.
1.50	UP TO	1.75	77.
1.75	UP TO	2.00	32.
2.00	UP TO	2.25	11.
2.25	UP TO	2.50	4.
2.50	UP TO	2.75	1.
2.75	UP TO	3.00	0.
3.00	AND OVER		0.

A second set of example data will further demonstrate the data description capabilities of the Pearson distributions. Suppose that the data in Table 3 represent grouped data from an unknown distributional form. The trial and error approach of interval selection and frequency distribution determination might produce the descriptions shown in Table 3 and Figure 4 or Table 4 and Figure 5. Once again the optimum or correct distributional form is left somewhat to conjecture. After computing the first four moments for this data the Pearson distributions produce the curve shown in Figure 6 with equation,

$$f(x) = \mathrm{Exp}\left[\log(.7629) + 1.680\ \log(1 + \frac{x}{.9975}) + 4.862\ \log(1 - \frac{x}{2.182})\right] \tag{4}$$

According to the program information supplied, the curve in Figure 6 is of the Pearson Type VI family. In actuality, the data in Table 3 and Table 4 were produced using a Weibul data generator. Since the Weibul distribution is contained in the Pearson Type VI, the efficacy of the description is again verified.

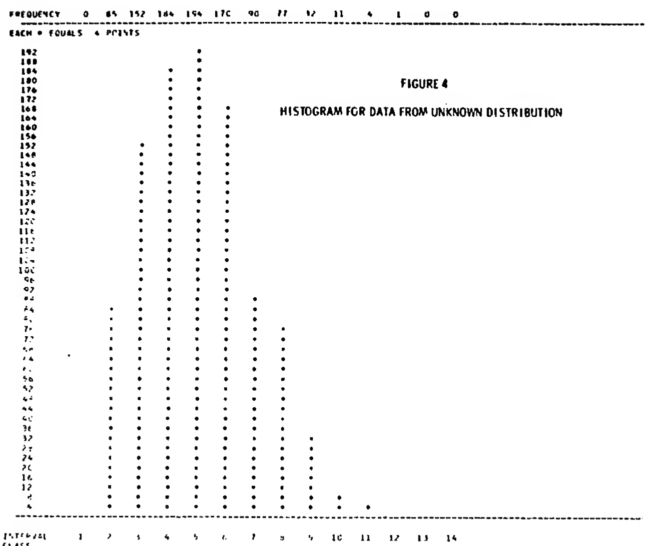


FIGURE 4
HISTOGRAM FOR DATA FROM UNKNOWN DISTRIBUTION

TABLE 4

Feasible Grouping for Unknown Distribution

	CLASS		FREQUENCY
	UNDER	0.0	0.
0.0	UP TO	0.13	22.
0.13	UP TO	0.25	63.
0.25	UP TO	0.38	69.
0.38	UP TO	0.50	83.
0.50	UP TO	0.63	80.
0.63	UP TO	0.75	104.
0.75	UP TO	0.88	97.
0.88	UP TO	1.00	97.
1.00	UP TO	1.13	86.
1.13	UP TO	1.25	84.
1.25	UP TO	1.38	55.
1.38	UP TO	1.50	35.
1.50	UP TO	1.63	42.
1.63	UP TO	1.75	35.
1.75	UP TO	1.88	18.
1.88	UP TO	2.00	14.
2.00	UP TO	2.13	6.
2.13	UP TO	2.25	5.
2.25	AND OVER		5.

As a final example, consider the data in Table 5 as representing grouped failure data drawn from an unknown distribution. Table 5 and Figure 7 or Table 6 and Figure 8 might represent reasonable data classifications and frequency distributions for the data as seen through the eyes of two different analysts. Figure 9 illustrates the Pearson description of the curve with equation,

$$f(x) = \mathrm{Exp}\left[\log(1.871) + 1.8471\ \log(1 + \frac{x}{.4833}) + 1.948\ \log\ (1 - \frac{x}{.5005})\right] \tag{5}$$

This curve is of the Pearson Type I.

In actuality the data in Table 5 and Table 6 were generated from a beta data generator. Since the beta distributions are contained in the Pearson Type I family, the efficacy of the description is once again verifed.

Experimentation with data selected from several distributions all indicated that the Pearson family of curves can be a powerful and efficient method for describing data drawn from unknown distributions. These other distributions from which data were generated and described included the normal, log-normal, and gamma distributions. In addition, a range of parametric values for each distribution was used so that a variety of curve shapes would manifest itself. In every case the Pearson descriptions provided immediate, accurate, and efficient plots for the data, as well as computations of the first four moments for each set of data and equations for each curve.

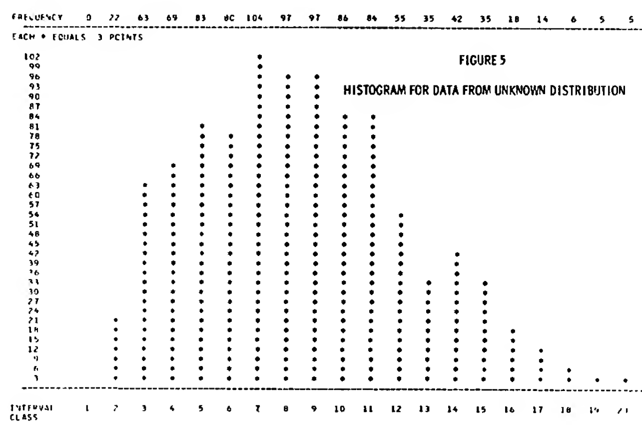


FIGURE 5
HISTOGRAM FOR DATA FROM UNKNOWN DISTRIBUTION

FIGURE 6 PEARSONIAN DESCRIPTION OF DATA FROM UNKNOWN DISTRIBUTION

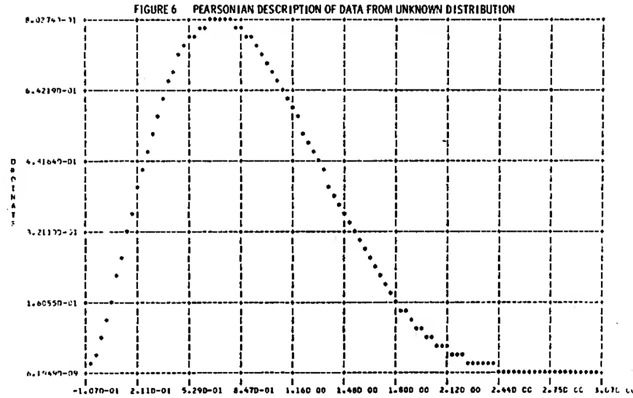


TABLE 5

Feasible Grouping for Unknown Distribution

CLASS			FREQUENCY
	UNDER	0.0	0.
0.0	UP TO	0.10	7.
0.10	UP TO	0.20	52.
0.20	UP TO	0.30	104.
0.30	UP TO	0.40	160.
0.40	UP TO	0.50	181.
0.50	UP TO	0.60	188.
0.60	UP TO	0.70	146.
0.70	UP TO	0.80	103.
0.80	UP TO	0.90	50.
0.90	UP TO	1.00	9.
1.00	AND OVER		0.

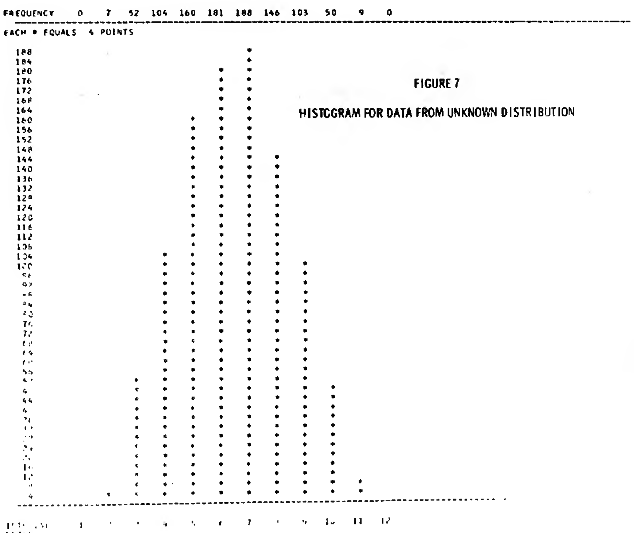


FIGURE 7

HISTOGRAM FOR DATA FROM UNKNOWN DISTRIBUTION

TABLE 6

Feasible Grouping for Unknown Distribution

CLASS			FREQUENCY
	UNDER	0.05	1.
0.05	UP TO	0.10	6.
0.10	UP TO	0.15	20.
0.15	UP TO	0.20	32.
0.20	UP TO	0.25	46.
0.25	UP TO	0.30	58.
0.30	UP TO	0.35	77.
0.35	UP TO	0.40	83.
0.40	UP TO	0.45	92.
0.45	UP TO	0.50	89.
0.50	UP TO	0.55	100.
0.55	UP TO	0.60	88.
0.60	UP TO	0.65	75.
0.65	UP TO	0.70	71.
0.70	UP TO	0.75	60.
0.75	UP TO	0.80	43.
0.80	UP TO	0.85	35.
0.85	UP TO	0.90	15.
0.90	UP TO	0.95	8.
0.95	AND OVER		1.

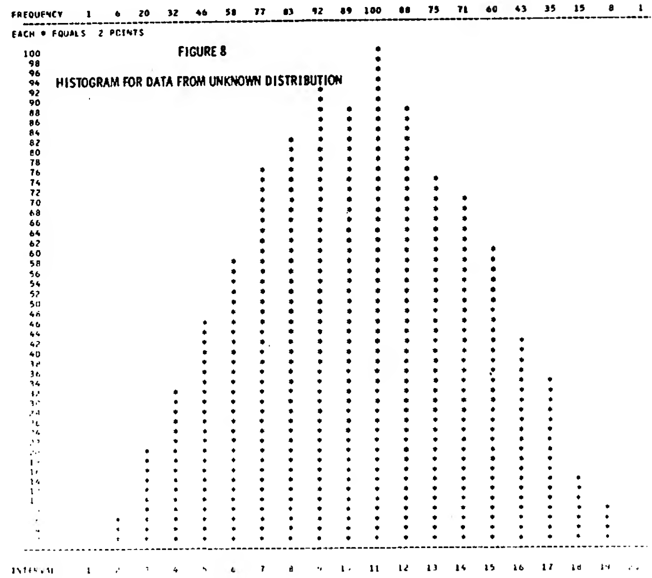
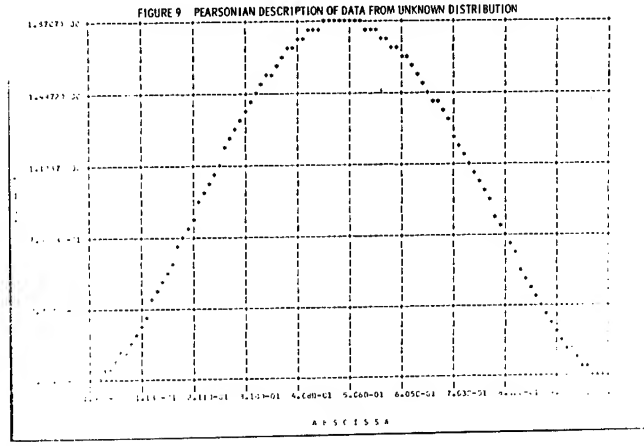


FIGURE 8

HISTOGRAM FOR DATA FROM UNKNOWN DISTRIBUTION

FIGURE 9 PEARSONIAN DESCRIPTION OF DATA FROM UNKNOWN DISTRIBUTION



Conclusion

When failure data are drawn from unknown distributions the typical trial and error procedures may not produce accurate descriptions of the data. Several generalized methods of data description are evident in the literature, including the Pearson curves. Although the Pearson distributions have sufficeint flexibility to include all known curve shapes they have not proven popular because of the mathematical expertise required. The Howe-Van Horn Pearson Data Describer provides a generalized method for efficiently and accurately describing data and eliminates the subjectivity inherent in probability plotting and goodness of fit tests. Experimentation with data drawn from several different distributional forms demonstrates the efficacy and power of the computer program and the broad descriptive properties of the Pearson curves. Using the program, a Pearson curve will be formulated for any set of data. The descriptive curve will be determined by the first four moments of the data and will provide a precise fit with certainty.

References

1. Bazovsky, Igor, Reliability: Theory and Practice, Prentice Hall, Englewood Cliffs, N.J., 1961.
2. Elderton, W. Palin, Frequency Curves and Correlation, Harren Press, Washington, D. C., 1953.
3. Hahn, G. K., and Shapiro, S. S., Statistical Models in Engineering, Wiley, New York, 1967.
4. Kendall, M. G., and Stuart, A., The Advanced Theory of Statistics, Vol. 1, 3rd Ed., Hofner Publishing Co., New York, 1947.
5. Pieruschka, Erich, Principles of Reliability, Prentise Hall, Englewood Cliffs, N. J., 1963.

SAMPLING PROGRAMS FOR RELIABILITY

Gerald J. Lieberman
Stanford University, Stanford, California 94305

0. Summary

This paper is concerned with six different types of programs to determine lower confidence bounds on $R(t_0)$, the reliability at time t_0, and on the time t_R, corresponding to a fixed reliability R. The six plans considered are various combinations of sampling with and without replacement, censored sampling, truncated sampling, and mixed censored-truncated sampling. Complete results are presented when the underlying time to failure distribution is assumed to be exponential, and partial results are given for the Weibull.

1. Introduction

All too frequently, the statistician fails the reliability engineer by giving him a solution to a problem he (the statistician) can solve, rather than a solution to the problem confronting the engineer. Examples of this phenomena in reliability occur when the failure time distribution is assumed to be exponential, and the method of testing is modified to suit the required assumptions. The problem of determining the form of the underlying distribution is one that has been treated in the literature, although not to the extent that it deserves, and will not be discussed further in this paper. The problem of choosing a testing program is one that deserves consideration and will be addressed.

Reliability of a component (or system) can be defined as the probability that it will perform satisfactorily for a specified period, t, in a given environment. In terms of the random variable, time to failure, henceforth denoted by X, this reliability can be expressed as

$$R(t) = P\{X > t\} = 1-F_X(t),$$

where $F_X(t)$ is the cumulative distribution function (CDF) of the random variable, time to failure. Determining this reliability becomes a major statistical problem in that estimates are required from experimental data. In order to obtain these data, testing programs must be initiated. A characterization of a testing program should include at least such information as (1) the number of units to be tested, (2) whether or not failed units will be replaced, and (3) how the test is to be terminated. Another criteria is the frequency with which the tested unit is checked, e.g., continuously or at fixed intervals, but it can usually be assumed that if checking is done at fixed intervals they are sufficiently narrow that continuous checking is a good approximation.

2. Testing Programs

The foregoing three criteria for determining a sampling plan will be used. Denote the number of units to be tested by n. The letter w will represent sampling with replacement and $\bar{w}$ will represent sampling without replacement. When sampling with replacement, it will be assumed that units that fail will be discovered immediately and instantaneously replaced by a new unit. The letter r will signify that the test is terminated at the time of the r^{th} failure, while the letter τ will signify that the test is terminated after a period of time τ has elapsed. Thus, there exist four types of sampling plans, namely

$[n=n_0,w,r=r_0]$: n_0 units are placed on test, failed units are replaced, and the test is terminated at the r_0^{th} failure.

$[n=n_0,w,\tau=\tau_0]$: n_0 units are placed on test, failed units are replaced, and the test is terminated after τ_0 hours.

$[n=n_0,\bar{w},r=r_0]$: n_0 units are placed on test, failed units are not replaced, and the test is terminated at the r_0^{th} failure.

$[n=n_0,\bar{w},\tau=\tau_0]$: n_0 units are placed on test, failed units are not replaced, and the test is terminated after τ_0 hours.

The plans $[n=n_0,w,r=r_0]$ and $[n=n_0,\bar{w},r=r_0]$ are often referred to as censored sampling plans. Precisely the data are said to be subjected to Type II censoring at r_0 out of n_0. The plans $[n=n_0,w,\tau=\tau_0]$ and $[n=n_0,\bar{w},\tau=\tau_0]$ are often referred to as truncated plans.

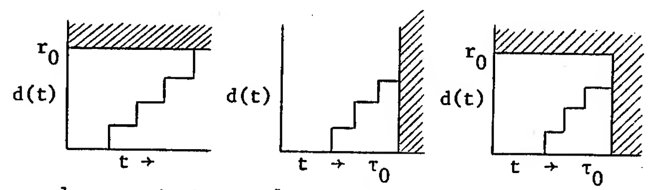
The aforementioned plans may sometimes have some practical difficulties associated with them. The censored plans may require testing for a long period of time before the r^{th} failure is observed. This would argue for using truncated plans, but this may be inefficient if many units fail quickly. Hence, some form of mixture of these two types of plans may be appropriate. In this mixed censored-truncated plan, it is decided in advance that the test is terminated at the time of the r_0^{th} failure, $X_{(r_0)}$, if $X_{(r_0)} < \tau_0$, and the test is terminated at time τ_0, if $X_{(r_0)} \geq \tau_0$, i.e., the test is terminated at $\min(X_{(r_0)},\tau_0)$. The notation for this type of plan will be given by

$[n=n_0,w,\min(X_{(r_0)},\tau_0]$: n_0 units are placed on test, failed units are replaced, and the test is terminated at the minimum of $X_{(r_0)}$ and τ_0.

$[n=n_0,\bar{w},\min(X_{(r_0)},\tau_0)]$: n_0 units are placed on test, failed units are not replaced, and the test is terminated at the minimum of $X_{(r_0)}$ and τ_0.

The notation $[n=100,w,\min(X_{(6)},150)]$ implies that the sampling plan will provide for placing 100 units on test with failed units being replaced. The test is terminated at the time of the 6^{th} failure, if this time is less than 150 hours, and is terminated at 150 hours if the 6^{th} failure has not yet occurred.

There exists an interesting pictorial representation of these sampling plans. Let d(t) denote the number of failures that occur during time t. A plot of d(t) versus t yields interesting termination regions (Figure 1).



a] censored plan b] truncated plan c] mixed censored-truncated plan

Figure 1 – Shaded Areas Indicate Termination Regions

Figure 1[a] shows a censored plan and the shaded area indicates that the test terminates when the r_0 failure occurs. Similarly, Figure 1[b] shows a truncated plan and the shaded area indicates that the test terminates after τ_0 hours. Finally, Figure 1[c] shows a mixed censored-truncated plan and the shaded area indicates that the test is terminated at the time of the r_0^{th} failure, if this time is less than τ_0 hours, and is terminated at τ_0 hours if the r_0^{th} failure has not occurred by this time. If should be noted that new types of plans can be generated by varying the shape of the termination regions. The usual Wald sequential analysis plan leads to parallel lines.

Given a testing Program, it still remains to "determine" the reliability. A usual experimental objective is to obtain a lower confidence bound for the reliability, $R(t_0)$, for some specified time t_0. The remainder of this paper will be devoted to this problem.

3. Testing Programs When the Time to Failure Has an Exponential Distribution

A very common assumption about the CDF for the random variable, time to failure, is that it is exponential, i.e.,

$$P\{X \le t\} = F_X(t) = \begin{cases} 1 - e^{-t/\theta}, & t \ge 0 \\ 0, & t < 0 \end{cases} \tag{1}$$

for $\theta > 0$.

The reliability at time t_0 is then given by $R(t_0) = e^{-t_0/\theta}$. There is extensive literature for finding lower confidence bounds for $R(t_0)$ based upon the exponential distribution. It is worth examining some of these results in the light of the testing programs previously introduced.

3.1 Sampling with Replacement - Censored w Plans

The mathematical results for plans that require replacement of units are "nice", in that they lead to exact confidence statements. Before presenting the results some notation will be introduced which will be required in all of the ensuing sections. Let $X_{(1)}$ be the random variable, time to the first failure, $X_{(2)}$ be the random variable, time to the second failure,..., and $X_{(r)}$ be the random variable, the time to the r^{th} failure. Let D be the random variable, the number of items that fail before time τ_0. Finally, denote by $\chi^2_{\nu;\alpha}$ the upper 100α percent point of the chi square distribution with ν degrees of freedom. These values are tabulated and may be found in any standard statistical textbook [1].

For the $[n=n_0,w,r=r_0]$ censored testing program, the 100γ percent lower confidence bound on $R(t_0)$ can be expressed as

$$\exp-\{t_0\, \chi^2_{2r_0;1-\gamma}/2n\, X_{(r_0)}\}. \tag{2}$$

Note that based upon the sample data, this is just a function of $nX_{(r_0)}$ which is just the total lifetime of all the units on test during the duration of the testing program. Since replacements are assumed to be made instantaneously there are always n units on test and the test terminates at the time of the r_0^{th} failure, i.e., $X_{(r_0)}$.

Example 1: Suppose 100 units are tested and the test is terminated at the time of the 5th failure which occurs at 910.7 hours. Find a 95% lower confidence bound on the reliability of the unit at $t_0 = 100$ hours, i.e., R(100).

Solution: Using expression (2) and noting that $\chi^2_{10,05} = 18.307$, the lower confidence bound is given by

$$\exp-\{(100)(18.307)/2(100)(910.7)\} = .990.$$

In the aforementioned example, the time t_0 is fixed beforehand one wishes to estimate the reliability corresponding to this time t_0. Suppose this is reversed and the reliability is fixed beforehand, and one wishes to obtain a lower confidence bound estimate of the time, t_R, corresponding to this reliability (note that t_R is defined by $P\{X > t_R\} = R$).

A 100γ percent lower confidence bound on t_R is given by

$$[2nX_{(r_0)} \log(1/R)]/\chi^2_{2r_0;1-\gamma}. \tag{3}$$

Example 2: Using the data in example 1, find a 95% lower confidence bound on the time, $t_{.90}$, corresponding to the reliability of .90.

Solution: Using expression (3), the lower confidence bound is given by

$$[2(100)\ 910.7 \log(1/.90)]/18.307 = 1048.$$

Expression (3) can be used to find the minimum time for the 5th failure to have occurred so as to be 95% confident that the reliability at time 100, i.e., R(100), is not less than .9. This is found by setting expression (3) equal to t_0 and solving for $X_{(5)}$, i.e.

$$X_{(5)} \ge \chi^2_{2r_0;1-\gamma}\, t_0/2n \log(1/R) =$$

$$18.307(100)/2(100)\log(1/.90) = 86.878.$$

3.2 Sampling with Replacement-Truncated w Plans

For the $[n=n_0,w,\tau=\tau_0]$ truncated testing program, the 100γ percent lower confidence bound on $R(t_0)$ can be expressed as

$$\exp-\{(t_0\chi^2_{2(D+1);1-\gamma})/2n\tau_0\}, \tag{4}$$

where D is the number of items that fail before τ_0. Note that based upon the sample data this is just a function of the number of items that fail before τ_0. Furthermore, D is a sufficient statistic so that the instants of failure $X_{(1)}, X_{(2)},\ldots,X_{(D)}$ contain no additional information about the reliability.

Example 3: Suppose 100 units are tested and the test is terminated after 600 hours, with 5 failures occuring. Find a 95% lower confidence bound on the reliability of a unit at $t_0 = 150$ hours, i.e., R(150).

Solution: Using expression (4) and noting that $\chi^2_{12;.05} = 21.026$, the lower confidence bound is given by

$$\exp-\{(150)(21.026)/2(100)(600)\} = .974.$$

Results for obtaining a lower confidence bound estimate of the time, t_R, corresponding to a fixed reliability R are given in Table 1.

3.3 Sampling with Replacement-Mixed Censored-Truncated w Plans

For the $[n=n_0,w,\min(X_{(r_0)},\tau_0)]$ mixed censored-

truncated testing program, the 100 percent lower confidence bound on $R(t_0)$ can be expressed as

$$\exp-\{(t_0\ \chi^2_{2(D+1);1-\gamma})/2n\tau_0\},\quad \text{for } X_{(r_0)} > \tau_0 \tag{5a}$$

and

$$\exp\{(t_0\ \chi^2_{2r_0;1-\gamma})/2n\ X_{(r_0)}\},\ \text{for } X_{(r_0)} \le \tau_0.$$

In this testing program, testing is terminated at either time τ_0 if the number of failures at that time is less than r_0 or at $X_{(r_0)}$, the time of the r_0^{th} failure provided that $X_{(r_0)}<\tau_0$. It is interesting to note that in this mixed censored-truncated program one uses the results for the $[n=n_0,w,\tau_0]$ plan if $X_{(r_0)} > \tau_0$ (expression 4) or the results from the $[n=n_0,w,r=r_0]$ plan if $X_{(r_0)} \le \tau_0$ (expression 2).

Example 4: Suppose 100 units are tested and the test is terminated after 600 hours or after the time of the 6th failure, whichever occurs first. Suppose that only 5 failures occurred at 600 hours. Find a 95% lower confidence bound on the reliability of a unit at t_0 = 150 hours, i.e., R(150).

Solution: Since less than 6 failures occurred by 600 hours, expression (5a) is relevant and the solution is the same as for example 3.

Results for obtaining a lower confidence bound estimate of the time, t_R, corresponding to a fixed reliability R are given in Table 1.

3.4 Sampling Without Replacement - Censored $\bar{w}$ Plans

The mathematical results for plans that require sampling without replacement generally are cumbersome except for the case of censored $\bar{w}$ plans. For the $[n=n_0,\bar{w},r=r_0]$ censored testing program, the 100γ percent lower confidence bound on $R(t_0)$ can be expressed as

$$\exp-\{(t_0\ \chi^2_{2r_0;1-\gamma})/2r_0\hat{\theta}\}, \tag{6}$$

where $r_0\hat{\theta} = \sum_{i=1}^{r_0} X_{(i)} + (n-r_0)X_{(r_0)}$ is just the total lifetime of all the units on test for the duration of the testing program. With this interpretation for $r_0\hat{\theta}$ expression (6) is the same as expression (2).

Example 5: A sample of size n=10 is taken and the test is terminated after the r=5th failure. The ordered failure times are as follows:

$X_{(1)}=50$, $X_{(2)}=75$, $X_{(3)}=125$, $X_{(4)}=250$ and $X_{(5)}=300$.

Find a 90% lower confidence bound for the reliability at t_0 = 40.

Solution: Using expression (6) and noting that $5\hat{\theta}$ = 50+75+125+250+300+(5)(300) = 2300, and $\chi^2_{10;.10}$ = 15.987, the lower confidence bound is given by

$$\exp-\{(40)(15.987)/(2)(2300)\} = .870.$$

If the reliability is fixed beforehand, and one wishes to obtain a lower confidence bound estimate of the time, t_r, corresponding to this reliability, the expression

$$[2r_0\hat{\theta}\ \log(1/R)]/\chi^2_{2r_0;1-\gamma} \tag{7}$$

is applicable, where $\hat{\theta}$ is defined as before.

Example 6: Using the date of example 5, find a 90% lower confidence bound for t_R corresponding to a reliability R = .85.

Solution: Substituting into expression (7) yields

$$[(2)(2300)\ \log(1/.85)]/15.987 = 46.762.$$

3.5 Sampling Without Replacement-Truncated $\bar{w}$ Plans

In order to obtain the "best" results for the $[n=n_0,\bar{w},\tau=\tau_0]$ plan, it is necessary to determine a lower confidence bound as a function of the number of units, D, that fail by time τ_0 and the total time on test, $T(\tau_0)$, where

$$T(\tau_0) = \sum_{i=1}^{D} X_{(i)} + (n-D)\tau_0.$$

Although in principle, such a bound can be obtained, in fact, the mathematics is very cumbersome and hence, two alternatives will be considered: [1] a lower confidence bound based solely on D and [2] an approximate result based upon both D and $T(\tau_0)$. The first alternative is attractive in that it leads to non-parametric results which will be useful for underlying distributions other than the exponential. However, it is inefficient when n is small since the information thrown away is important. This non-parametric 100γ percent lower confidence bound is given by

$$\left[1+\left(\frac{D+1}{n-D}\right) F_{2(D+1),\ 2(n-D);1-\gamma}\right]^{-\frac{t_0}{\tau_0}} \tag{8}$$

where $F_{\nu_1\nu_2,\alpha}$ is the upper 100α percent point of the F distribution with ν_1 and ν_2 degrees of freedom. The approximate 100γ percent lower confidence bound is given by

$$\exp-\{(t_0\ \chi^2_{2(D+1);1-\gamma})/2T(\tau_0)\},\quad \text{for } D=0,1,\ldots,n-1$$

and (9)

$$\exp-\{(t_0\ \chi^2_{2n;1-\gamma})/2T(X_{(n)})\},\quad \text{for } D=n,$$

where

$$T(Z) = \sum_{i=1}^{D} X_{(i)} + (n-D)Z.$$

Example 7: 20 units are placed on test and the test is terminated at 100 hours. Two items failed at 80 and 93 hours, respectively. Find a 95% lower confidence bound for the reliability at time 100 hours, i.e., R(100).

Solution: The non-parametric result will be obtained first. Noting that $F_{6,36;05}$ = 2.36 and substituting into (8), the lower confidence bound is given by

$$\left[1+\left(\frac{3}{18}\right)2.36\right]^{-1} = .718.$$

The approximate 95% lower confidence bound will be obtained next. Noting that

$$\chi^2_{6;05} = 12.592, \quad\text{and}$$

$$T(100) = 80 + 93 + 18(100) = 1973,$$

and substituting into expression (9), the 95% approximate

lower confidence bound is given by

$$\exp-\{(100)12.592/(2)(1973)\} = .727.$$

Note the closeness of these two results, even for relatively small n.

Results for obtaining a lower confidence bound estimate of the time, t_R, corresponding to a fixed reliability R are given in Table 1.

3.6 Sampling with Replacement - Mixed Censored Truncated $\bar{w}$ Plans

For the $[n=n_0,\bar{w},\min(X_{(r_0)},\tau_0)]$ mixed censored-truncated testing program, the exact results are cumbersome. Approximate 100γ percent lower confidence bounds are given by

$$\exp-\{(t_0\chi^2_{2(D+1);1-\gamma})/2T(\tau_0), \text{ when } X_{(r_0)} > \tau_0 \qquad (10a)$$

and

$$\exp-\{(t_0\chi^2_{2r_0;1-\gamma})/2r_0\hat{\theta}, \qquad \text{when } X_{(r_0)} \le \tau_0 \qquad (10b)$$

where $$T(\tau_0) = \sum_{i=1}^{D} X_{(i)} + (n_0-D)\tau_0$$

and $$r_0\hat{\theta} = \sum_{i=1}^{r_0} X_{(i)} + (n-r_0)\tau_0.$$

Note the similarity with the expressions for the $[n=n_0,\bar{w},r=r_0]$ (expression 6) and the $[n=n_0,\bar{w},\tau=\tau_0]$ (expression 9).

Example 8: Suppose 10 units are tested and the test is terminated after 400 hours or after the time of the 5th failure, whichever occurs first. Suppose that the failure times are as follows:

$X_{(1)}=50$, $X_{(2)}=75$, $X_{(3)}=125$, $X_{(4)}=250$ and $X_{(5)}=300$.

Find a 90% lower confidence bound for the reliability at $t_0 = 40$.

Solution: Since the 5th failure occurred at 300 hours, expression (10b) is appropriate and the example is similar to example 5. The total time on test $5\hat{\theta} = 50 + 75 + 125 + 250 + 300 + 5(400) = 2800$ so that the lower confidence bound is given by

$$\exp-\{(40)(15.987)/(2)(2800)\} = .892.$$

4. Testing Programs when the Time to Failure has a Distribution other than the Exponential

The problems that occur with the use of the exponential distribution have been widely discussed. In particular, the exponential distribution is a one parameter class which has the property of having a constant instantaneous failure rate. This unrealistic assumption has led researchers to consider other families of distributions, e.g., the Weibull, gamma, log normal, etc.

Perhaps the most important alternative to the exponential is the Weibull distribution. The cumulative distribution function for a random variable X having a two parameter Weibull distribution is given by

$$F_X(t) = \begin{cases} 1- e^{-(t/\alpha)^\beta}, & t \ge 0 \\ 0 & , t < 0 \end{cases} \qquad (11)$$

for $\alpha,\beta > 0$.

The parameter α is referred to as the scale parameter and β is called the shape parameter. The reliability, $R(t_0)$, at time t_0 is given by

$$R(t_0) = e^{-(t_0/\alpha)^\beta}.$$

Note that if β is set equal to 1, expressions (11) and (1) become identical so that the Weibull and exponential coincide, i.e., the exponential is just a special case of the Weibull. Furthermore, the instantaneous failure rate is a monotonically increasing function of t for $\beta > 1$ (aging) and a monotonically decreasing function of t for $\beta < 1$.

As an indication of the state of knowledge about testing programs when the time to failure has a Weibull distribution, it should be noted that no exact results exist when sampling is done with replacement, i.e., $\bar{w}$ plans. There are exact results for $[n=n_0,\bar{w},r=r_0]$ censored plans and they appear in a paper by Johns and Lieberman [4]. Unfortunately, simple expressions are unattainable and tables are required. The Johns-Lieberman paper presents tables of exact lower confidence bounds for the reliability for sample sizes (n) of 10, 15, 20, 30, 50, and 100, and for various values of r_0 and confidence coefficients γ. These tables can also be used to get lower confidence bounds on the time, t_R, corresponding to a fixed reliability R.

The non-parametric results contained in Section 3.5 are applicable to $[n=n_0,\bar{w},\tau=\tau_0]$ truncated plans. In particular, expression (8) leads to lower confidence bounds, but again is "efficient" only if n_0 is relatively large.

There are not exact results for the mixed censored-truncated plan $[n=n_0,\bar{w},\min(X_{(r_0)},\tau_0)]$ but an interesting conjecture is to use expression (8) when $X_{(r_0)} > \tau_0$ and use the Johns-Lieberman results when $X_{(r_0)} \le \tau_0$.

Results for other than the Weibull two parameter families of distributions are sparse. Exact parametric results for any of the six sampling programs using the two parameter gamma distribution are unknown. When one of the parameters are assumed to be known, results are obtainable. However, this is an unrealistic assumption, and furthermore, such an assumption is equivalent to assuming an exponential time to failure, so that these results become applicable.

REFERENCES

[1] Bowker, A.H. and Lieberman, G.J., Engineering Statistics, 2nd Edition, Prentice Hall, Englewood Cliffs, New Jersey, 1972.

[2] Department of Commerce, Office of Technical Services, Statistical Techniques in Life Testing, PB 171 580.

[3] Gnedenko, B.V., Belyayev, Y.K. and Solovyev, A.D., Mathematical Methods of Reliability Theory, Academic Press, New York and London, 1969.

[4] Johns, M.V. and Lieberman, G.J., "An Exact Asymptotically Efficient Confidence Bound for Reliability in the Case of the Weibull Distribution", Technometrics, Vol. 8, No. 1, pp. 135-175, February 1966.

Table 1

Formulas for Confidence Interval Estimators of the Reliability Based upon an Underlying Exponential Distribution

Testing Program	100γ Percent Lower Confidence Bound on $R(t_0)$	100γ Percent Lower Confidence Bound on a Quantile t_R, where t_R is that life such that $P\{X \geq t_R\}=R$.	Remarks
$[n=n_0, w, r=r_0]$	$\exp-\{(t_0\chi^2_{2r_0;1-\gamma})/2nX_{(r_0)}\}$	$[2nX_{(r_0)}\log(1/R)]/\chi^2_{2r_0;1-\gamma}$	
$[n=n_0, \bar{w}, r=r_0]$	$\exp-\{(t_0\chi^2_{2r_0;1-\gamma})/2r_0\hat{\theta}\}$	$[2r_0\hat{\theta}\log(1/R)]/\chi^2_{2r_0;1-\gamma}$	$r_0\hat{\theta}=\sum_{i=1}^{r_0} X_{(i)}+(n-r_0)X_{(r_0)}$
$[n=n_0, w, \tau=\tau_0]$	$\exp-\{(t_0\chi^2_{2(D+1);1-\gamma})/2n\tau_0\}$	$[2n\tau_0\log(1/R)]/\chi^2_{2(D+1);1-\gamma}$	D=number of items which fail before τ_0.
$[n=n_0, \bar{w}, \tau=\tau_0]$	$\left[1+\left(\frac{D+1}{n-D}\right)F_{2(D+1),2(n-D);1-\gamma}\right]^{-\frac{t_0}{\tau_0}}$ (nonparametric result) $\exp-\{(t_0\chi^2_{2(D+1);1-\gamma})/2T(\tau_0)\}$, for $D=0,1,\ldots,n-1$ $\exp-\{(t_0\chi^2_{2n;1-\gamma})/2T(X_{(n)})\}$, for $D=n$ (approximate parametric result)	$2T(\tau_0)\log(1/R)/\chi^2_{2(D+1);1-\gamma}$, for $D=0,1,2,\ldots,n-1$ $2T(X_{(n)})\log(1/R)/\chi^2_{2n;1-\gamma}$, for $D=n$	D=number of items which fail before τ_0. $T(Z) = \sum_{i=1}^{D} X_{(i)}+(n-D)Z$
$[n=n_0, w, \min(X_{(r_0)}, \tau_0)]$	$\exp-\{(t_0\chi^2_{2(D+1);1-\gamma})/2n\tau_0\}$, when $X_{(r_0)} > \tau_0$ $\exp-\{(t_0\chi^2_{2r_0;1-\gamma})/2nX_{(r_0)}\}$, when $X_{(r_0)} \leq \tau_0$	$[2n\tau_0\log(1/R)]/\chi^2_{2(D+1);1-\gamma}$, when $X_{(r_0)} > \tau_0$ $[2r_0\hat{\theta}\log(1/R)]/\chi^2_{2r_0;1-\gamma}$, when $X_{(r_0)} \leq \tau_0$	D=number of items which fail before τ_0.
$[n=n_0, \bar{w}, \min(X_{(r_0)}, \tau_0)]$	$\exp-\{(t_0\chi^2_{2(D+1);1-\gamma})/2T(\tau_0)\}$, when $X_{(r_0)} > \tau_0$ $\exp-\{(t_0\chi^2_{2r_0;1-\gamma})/2r_0\hat{\theta}\}$, when $X_{(r_0)} \leq \tau_0$ (approximate parametric result)	$[2T(\tau_0)\log(1/R)]/\chi^2_{2(D+1);1-\gamma}$, when $X_{(r_0)} > \tau_0$ $[2r_0\hat{\theta}\log(1/R)]/\chi^2_{2r_0;1-\gamma}$, when $X_{(r_0)} \leq \tau_0$	$T(\tau_0) = \sum_{i=1}^{D} X_{(i)}+(n-D)\tau_0$, D=number of items which fail before τ_0, $r_0\hat{\theta} = \sum_{i=1}^{r_0} X_{(i)}+(n-r_0)\tau_0$

CENSORED SAMPLE SIZE SELECTION FOR LIFE TESTS

INDEX SERIAL NUMBER – 1079

John I. McCool
SKF Industries, Inc.
Engineering and Research Center
King of Prussia, Pennsylvania

Abstract

The proper choice of the sample size n and the number of failures r in designing a life test is shown to be governed by the use to which the test results will be put.

Examples are given of the differences in sample sizes that result when the purposes are:

1. The conduct of an acceptance test based on a specific percentile value of the life distribution.

2. The determination of a "historical" value of a percentile.

The examples are based on existing results for the exponential distribution and some more recent results for the Weibull distribution.

Introduction

Although guidance with sample size selection is a major reason that engineers will seek the support of a statistician, it is not a heavily emphasized topic in statistics curricula. In particular the fact that the proper choice of sample size is critically dependent on the purpose of the test is not heavily stressed.

The purpose of this paper is to illustrate and compare two methods of selecting the sample size and number of failures for a censored life test from which inferences are to be made regarding a specific percentile (or reliable life) of a time-to-failure distribution.

The type of censoring considered is r out of n, or Type II censoring, under which n items are tested until the first r fail. Time to failure is thus a random variable.

The two methods for sample size selection are evaluated for both the one parameter exponential and the two parameter Weibull models.

The first method is based on fixing two points on the operating characteristic curve of a hypothesis test and will be called the OC curve criterion.

The second method is based on fixing the ratio of the upper to lower ends of a two sided confidence interval and will be called the confidence interval criterion.

In order not to be distractingly general in what follows we will, for the most part, assume that the life test is being conducted for the purpose of making inferences about the tenth percentile, $X_{0.10}$, of the time to failure distribution. $X_{0.10}$ is the life which 90% of the items in the population will exceed and is therefore sometimes called the 90% reliable life. The principles, of course, are applicable to any percentile.

OC Curve Criterion

We consider the acceptance test situation where the experimenter's purpose is to determine by means of a life test whether $X_{0.10}$ for a batch of new material is as good as a standard or design value $(X_{0.10})_o$. He is only concerned with detecting when $X_{0.10}$ is less than $(X_{0.10})_o$. Having performed the life test he will use the results to compute an estimate $\hat{X}_{0.10}$ of $X_{0.10}$. If this estimate is sufficiently large he will be satisfied that $X_{0.10}$ is adequate for the material on which the test was performed. If $\hat{X}_{0.10}$ is small he will reject the claim that the new material has an $X_{0.10}$ value as good as the standard value $(X_{0.10})_o$.

Mathematically, he will determine a critical value C and,

$$\text{Accept if } \hat{X}_{0.10} > C$$

$$\text{Reject if } X_{0.10} < C$$

Because $\hat{X}_{0.10}$ is an estimate based on a finite sample, the experimenter realizes that $\hat{X}_{0.10}$ won't be equal to the "true" value $X_{0.10}$. The true value of $X_{0.10}$ is unknown and could be determined only by testing an indefinitely large sample.

He therefore realizes that by making a decision based on $X_{0.10}$, that that decision could be wrong.

Specifically he could err in the following two ways:

1) He could reject the material although $X_{0.10} \geq (X_{0.10})_o$ (because it turned out that $\hat{X}_{0.10} < C$)

2) He could accept the material although $X_{0.10} \leq (X_{0.10})_o$ (because it turned

out that $\hat{X}_{0.10} > C$)

By suitable choice of C, the experimenter can limit the probability of the first of these two types of errors. By suitable choice of sample size he can also limit the probability of committing the second type of error.

We will consider selection of C and sample size for the case where the experimenter is using the maximum likelihood estimate of $X_{0.10}$ and wherein the underlying distribution is either exponential or Weibull.

Exponential Distribution

a) OC Curve Criterion

The choice of C which will assure that the probability of rejecting is not greater than the small value $\alpha = 0.10$, if $X_{0.10} = (X_{0.10})_1 \leq (X_{0.10})_o$, is calculated using Eq. (A1-4) of Appendix I as,

$$C = \frac{(X_{0.10})_o \; \chi^2_{0.10}(2r)}{2r}$$

where $\chi^2_{0.10}(2r)$ is the tenth percentile of the χ^2 distribution with 2r degrees of freedom, and where r is the number of failures obtained in the life test. This value is independent of the sampe size n in the exponential case.

For example, if the historical or design value of $X_{0.10}$ was $(X_{0.10})_o = 1500$ hrs. and if r = 15 failures ($\chi^2_{0.10}(30) = 20.6$) were being considered, then the rule of action

$$\text{Accept if } \hat{X}_{0.10} > C = \frac{1500 \times 20.6}{30} = 1030 \text{ hrs.}$$

will assure that the probability of rejecting, if $X_{0.10}$ were truly ≥ 1500, would be 0.10 or less. Thus under this plan good material will not be rejected very often.

On the other hand using this procedure bad material will be accepted with a probability that depends on how bad it is. At this point the experimenter must face squarely the question of how bad is really bad. Obviously if $X_{0.10} = 1500$ hrs. is considered good he will not consider that $X_{0.10} =$ 1495 hrs. is terribly bad. On the other hand $X_{0.10} = 500$ hours might be acceptable or unacceptable depending upon the application. Using Eq. (A1-6) of Appendix I one can calculate the value of $(X_{0.10}) = (X_{0.10})_1$ below which mistaken acceptance will not occur with a probability greater than γ. Taking $\gamma = 0.10$, one calculates for the present example with $(X_{0.10})_o = 1500$ and r = 15,

$$(X_{0.10})_1 = (X_{0.10})_o \; \frac{\chi^2_{0.10}(30)}{\chi^2_{0.90}(30)}$$

$$= 1500 \times \frac{20.6}{40.3} = 770 \text{ hrs.}$$

Thus, with r = 15, any values of $X_{0.10} = (X_{0.10})_1$ below 770 hours will be mistakenly accepted with a probability less than 0.10.

This may or may not be adequate for the experimenter's purpose. Rather than picking r values and calculating $(X_{0.10})_1$ the experimenter should instead specify the values of $(X_{0.10})_1$ below which he wishes the probability of mistaken acceptance to be less than 0.10 (or any other small probability).

The solid curve in Figure 1 is a plot as a function of r of the natural logarithm of the ratio of $(X_{0.10})_o/(X_{0.10})_1 = \chi^2_{0.90}(2r)/\chi^2_{0.10}(2r)$ for a 10% probability of mistakenly rejecting when $X_{0.10} = (X_{0.10})_o$ and 10% probability of mistaken acceptance if $X_{0.10} = (X_{0.10})_1$. (The same curve in Figure 1 is actually applicable for a general percentile X_p rather than just for $X_{0.10}$).

Continuing with the foregoing example, consider finding the value of the number of failures r such that the probability of acceptance is 0.10 (or less) for $(X_{0.10})_1 =$ 900 hours.

In this case,

$$\log (X_{0.10})_o/(X_{0.10})_1 = \log \frac{1500}{900} = 0.51$$

From Figure 1 with $\log (X_{0.10})_o/(X_{0.10})_1 =$ .51, one finds r = 25. The value of C is then

$$C = 1500 \times \frac{37.7}{50} = 1130 \text{ hours}$$

Figure 2 illustrates the relationships involved. The bottom curve in Figure 2 is the exponential failure distribution when $X_{0.10} = 1500$ hours. The next curve up in Figure 2 is a plot on the same horizontal scale of the distribution of the ml estimate $\hat{X}_{0.10}$ that results when samples of size 25 are taken. It is seen that 90% of the area under this distribution is to the right of C = 1130 hours.

The third curve from the bottom in Figure 2 is the exponential time to failure distribution for which $X_{0.10} = 900$ hours. The associated distribution of $\hat{X}_{0.10}$ is shown as the topmost curve. On this curve the area to the right of C = 1130 hours is 10%.

Summarizing, we have selected a test under

which a sample of size n will be tested until the first 25 failures occur. The choice of n is, for the exponential distribution, up to the discretion of the experimenter. If the results are not needed quickly and material is expensive he should take $n = 25$.

Larger n should be used if material is cheap and the results are needed quickly. Having obtained the 25 failures he will calculate the ml estimate $X_{0.10}$ and accept if

$$\hat{X}_{0.10} \geq 1130 \text{ hours}$$

In so doing he is guaranteed that if $X_{0.10}$ is actually greater than 1500 hours, he won't reject it more often than 10% of the time. Conversely if $X_{0.10}$ is actually less than 900 hours, he won't mistakenly accept it more than 10% of the time.

Had a percentile other than $X_{0.10}$ been of concern the results would have been the same, i.e. if $(X_{.50})_0$ had been specified as 1500 hours, $\hat{X}_{0.50} \geq 1130$ would still have been the acceptance criterion and $(X_{0.50})_1 =$ 900 hours would have been the value at which the acceptance probability drops to 10%.

b) Confidence Interval Criterion

When the purpose of testing is not to test a hypothesis about a specific percentile but instead to estimate the value of the percentile as a part of a general information gathering program, the above considerations for sample size selection are no longer applicable. In this case the experimenter will use the results of the life test to calculate a point estimate as well as confidence limits for the true value of the percentile. It is reasonable in this case for him to select the sample size in such a way that he is sure to have determined the true value within a satisfactory degree of precision. A convenient and useful criterion is to set a value on the ratio R of the upper to lower ends of a two sided confidence interval.[4]

For the exponential distribution the ratio R for an 80% confidence interval is given by Eq. (A1-9) of Appendix I as,

$$R = \frac{\chi^2_{0.90}(2r)}{\chi^2_{0.10}(2r)}$$

Where r is the number of failures used in the life test.

Figure 3 is a plot of log R against number of failures r. Using this plot one may deduce the number of failures r, corresponding to a given value of the desired ratio R of the upper to lower ends of an 80% confidence interval.

It is noted that this expression relating R to r for an 80% confidence interval is exactly the same as the expression relating $(X_{0.10})_0/(X_{0.10})_1$ to r when the probabilities of the two types of error are both taken to be equal to 0.10.

Thus, for the exponential distribution, the OC curve criterion and the confidence interval criteria are fundamentally the same.

Weibull Distribution

a) OC Curve Criterion

If the underlying time to failure distribution is the two parameter Weibull with the shape parameter β unknown, the ml estimates of $X_{0.10}$ and β are obtained from the results of a Type II censored life test from Eqs. (A2-2) and (A2-3) of Appendix II.

Again $X_{0.10}$ will be accepted as being equal to $(X_{0.10})_0$ if

$$\hat{X}_{0.10} > C$$

If it is desirable that the probability of falsely rejecting good material be less than 0.10, one calculates C from Eq. (A2-5) of Appendix II as,

$$C = (X_{0.10})_0 \exp\left[\frac{u_{0.10}(r,n,0.10)}{\hat{\beta}}\right]$$

$u_{.10}(r,n,0.10)$ is the 10th percentile of the random variable $u(r,n,p = 0.10)$ defined in Appendix II. Unlike the exponential, a different value of C is needed in the Weibull case if a percentile other than $X_{0.10}$ is being investigated. For example, for the median $X_{0.50}$, C would be determined as,

$$C = (X_{0.50})_0 \exp\left[\frac{u_{0.10}(r,n,0.50)}{\hat{\beta}}\right]$$

The distributions $u(r,n,p)$ must be determined by Monte Carlo sampling. Percentage points for some r,n, and $p = 0.10$ are given in Table 1.

It is noted that C is not a fixed constant for the Weibull case but is a function of the random variable $\hat{\beta}$ estimated from the endurance test results.

The value $(X_{0.10})_1 < (X_{0.10})_0$ for which the probability of erroneous acceptance is 0.10 or less in the Weibull case is shown in Appendix II, to be given by

$$(X_{0.10})_1 = (X_{10})_0 \exp\left[-s_{0.10}(0.10,r,n,0.10)\right]$$

Where $s_{0.10}(0.10,r,n,0.10)$ is the tenth

percentile of a random variable $s(\alpha, r, n, p)$ for $\alpha = 0.10$, $p = 0.10$. The distribution of $s(\alpha, r, n, p)$ must be determined by Monte Carlo sampling.

$(X_{0.10})_1$ depends on the true but unknown value of the Weibull shape parameter. This is analogous to the situation in normal distribution theory that the OC curve for testing a hypothesis about the normal mean depends on the true but unknown standard deviation.

Figure 1 shows a plot of $\beta \log \left[(X_{0.10})_o/(X_{0.10})_1\right]$ against r for n = 10, 20 and 30. The values are obtained from Monte Carlo simulation of 2000 samples for each n and r. Some scatter is evident in the data points.

It is noted that for $r > 4$ the larger samples give superior discrimination (larger $(X_{0.10})_1$).

The plots become nearly horizontal as r increases indicating that there is relatively little to gain by waiting for additional failures after the curves flatten out. At r = 3 the curves have inverted and n = 10 gives superior discrimination to n = 30.

Since the Weibull distribution reduces to the exponential when $\beta = 1.0$, the curves of Figure 1 illustrate the loss in discrimination that results from assuming the Weibull and going through the estimation of β when the exponential model actually applies.

The values in Figure 1 are tabled as $\left[(X_{0.10})_o/(X_{0.10})_1\right]^\beta$ in Table 1. (For the exponential $\beta = 1.0$). Also shown in Table 1 are the values corresponding to 50% probability of accepting if $X_{0.10} = (X_{0.10})_1$.

As an example assume,

$$(X_{0.10})_o = 100 \text{ hours}$$

$$\beta = 2.0$$

Find n and r such that the probability of acceptance is 0.10 if $(X_{0.10})_1 = 55$ hours.

$$\left[X_{0.10}/(X_{0.10})_1\right]^\beta =$$

From Table 1 it is seen that either n = 20, r = 6 or n = 30, r = 6 should suffice. The experimenter may make the choice based on a trade-off between waiting time for results and the cost of test units.

If for example n = 20, r = 6 is chosen, the experimenter will, at the completion of the test, calculate $\hat{X}_{0.10}$ and $\hat{\beta}$ from the test results and accept if

$$\hat{X}_{0.10} > 100 . \exp\left(\frac{-0.61}{\hat{\beta}}\right)$$

wherein $u_{0.10}(6, 20, 0.10) = -0.61$ was obtained from Table I.

b) Confidence Interval Criterion

It is shown in Appendix II that the ratio of upper to lower confidence limits for a Weibull percentile is a random variable. The median ratio $R_{.50}$ is therefore proposed as a criterion for expressing the precision with which a Weibull percentile is determined by a test of n samples with r failures. As with the OC curve criterion, $R_{.50}$ depends on the true but unknown value of the Weibull shape parameter.

For 80% confidence limits on $X_{0.10}$, Eq. (A2-13) of Appendix II gives $R^{\beta}_{.50}$ as,

$$R^{\beta}_{0.50} = \exp\left[\frac{-u_{0.90}(r, n, p=0.10) + u_{0.10}(r, n, p=0.10)}{v_{0.50}(r, n)}\right]$$

Where $u_{0.90}$, $u_{0.10}$ and $v_{0.50}$ must be determined by Monte Carlo sampling.

Figure 3 is a plot of $\beta \log R_{0.50}$ thus computed against r for various n. Each data point on this plot was computed from a Monte Carlo experiment in which 10,000 samples were generated for each n and r.

The curves seem to grow flatter as sample size increases, are uniformly higher than the plot of log R for the exponential distribution, and do not overlap, at least for values of $r > 5$. Figure 4 is a similar plot for the case where the Weibull median $X_{0.50}$ is being estimated.

In this case the curves are much closer to the exponential curve, exhibit much more pronounced curvature and significant overlap. 5 failures in a sample of size 10 is seen to give a better estimate of $X_{0.50}$ than 5 failures in a larger sample. On the other hand, 20 failures in a group of size 30 is superior to a complete sample of size 20.

Table II lists the values of $R^{\beta}_{.50}$ for 80% confidence limits on $X_{.10}$ and $X_{.50}$ for the various sample sizes along with the value of R for the exponential case.

From this table it can be noted that n = 10, r = 3 is just about equally good (or bad) for the Weibull tenth percentile as for the Weibull median.

n = 30, r = 5 is better for $X_{0.10}$ than $X_{0.50}$ but n = 30, r = 10 is better for $X_{0.50}$ than $X_{0.10}$.

In general it appears that if more than 50% of the sample is failed, the median is

better estimated than $X_{0.10}$. If less than 50% are failed, $X_{0.10}$ may be better estimated than $X_{0.50}$.

As an example of the use of the confidence interval criterion for selecting sample size for a Weibull distribution, consider the case where the experimenter wishes that the median ratio of upper to lower 80% confidence limits for $X_{0.50}$ be 2.0 or less.

Assuming $\beta = 2$ gives,

$$\beta \log R_{.50} = 2 \times .693 = 1.39$$

From Figure 4 it is seen that any of the following choices suffice:

n = 10, r = 5
n = 15, r = 6
n = 20, r = 7
n = 30, r = 8

Closure

We have seen that for the Weibull distribution the appropriate number of failures depends upon:

1) The criterion used; OC curve or confidence interval ratio.

2) The percentile about which inferences are to be made.

3) The sample size.

For the exponential distribution, the number of failures indicated by the two criteria coincide if the probabilities of the two types of errors in the OC curve criterion are taken equal to each other and the confidence level of the confidence interval criterion is the complement of their sum. The appropriate number of failures is independent of sample size and the population being considered.

Appendix I

Exponential Distribution

a) OC Curve Criterion

Let X_p denote the 100 p-th percentile (or (1-p)th reliable life) of the one parameter exponential distribution. The cumulative distribution function (CDF) may be parameterized as,

$$F(x) = 1 - \exp - K_p\left(\frac{x}{X_p}\right) \qquad \text{(A1-1)}$$

where

$$K_p \equiv \log \frac{1}{1-p} \qquad \text{(A1-2)}$$

The maximum likelihood (ml) estimate of X_p is computed from the observations X_i in a sample of size n censored the r-th failure (Type II censoring), as,

$$\hat{X}_p = \frac{K_p}{r} \sum_{i=1}^{n} x_i \qquad \text{(A1-3)}$$

Further, $\frac{2r\hat{X}_p}{X_p}$ is distributed as $\chi^2(2r)$.

To test, at the 100α % level, the hypothesis that X_p assumes a specific value $(X_p)_o$, i.e. Ho: $X_p = (X_p)_o$ against the one sided alternative that X_p is less than $(X_p)_o$, i.e. H_1: $X_p = (X_p)_1 \leq (X_p)_o$.

One computes $\hat{X}_p$ and accepts Ho in favor of H_1 if,

$$\hat{X}_p > \frac{(X_p)_o \chi^2_\alpha(2r)}{2r} \qquad \text{(A1-4)}$$

where $\chi^2_\alpha(2r)$ denotes the 100α -th percentile of the χ^2 distribution having 2r degrees of freedom.

If Ho is true, i.e. if $X_p = (X_p)_o$, the probability of acceptance will be $1-\alpha$.

If H_0 is false, i.e. if $X_p = (X_p)_1 < (X_p)_o$ the probability of acceptance will be less than $1-\alpha$ by an amount that depends on the "falseness" of the null hypothesis, or more specifically, on the ratio $(X_p)_o/(X_p)_1$. This follows from

$$P_a = \text{Prob}\left[\hat{X}_p > \frac{(X_p)_o \chi^2_\alpha(2r)}{2r}\right] \qquad \text{(A1-5)}$$

$$= \text{Prob}\left[\frac{2r\hat{X}_p}{(X_p)_1} = \chi^2 > \frac{(X_p)_o}{(X_p)_1}\chi^2_\alpha(2r)\right]$$

Corresponding to Pa = γ, one has

$$\frac{(X_p)_o}{(X_p)_1}\chi^2_\alpha(2r) = \chi^2_{1-\gamma}(2r)$$

or

$$\frac{(X_p)_o}{(X_p)_1} = \frac{\chi^2_{1-\gamma}(2r)}{\chi^2_\alpha(2r)} \qquad \text{(A1-6)}$$

The operating characteristic curve is a plot of γ against $(X_p)_o/(X_p)_1$.

γ will decrease as $(X_p)_o/(X_p)_1$ decreases.

b) Confidence Interval Criterion

Two sided $100(1-\alpha)$% confidence intervals for X_p follow on inverting the inequalities in the statement

$$\text{Prob}\left[\chi^2_{\alpha/2}(2r) < \chi^2 = \frac{2r\hat{X}_p}{X_p} < \chi^2_{1-\alpha/2}(2r)\right] \qquad \text{(A1-7)}$$

They are,

$$\frac{2r\hat{X}_p}{\chi^2_{1-\alpha/2}(2r)} < X_p < \frac{2r\hat{X}_p}{\chi^2_{\alpha/2}(2r)} \quad \text{(A1-8)}$$

The ratio R of the upper to lower confidence limit is thus,

$$R = \frac{\chi^2_{1-\alpha/2}(2r)}{\chi^2_{\alpha/2}(2r)} \quad \text{(A1-9)}$$

Appendix II

Weibull Distribution

The CDF of the Weibull distribution is expressible in terms of the p-th percentile as,

$$F(x) = 1 - \exp - K_p\left(\frac{x}{x_p}\right)^\beta \quad \text{(A2-1)}$$

where β is the Weibull shape parameter.

The ml estimate of X_p is computed from a sample of size n, Type II censored at r failures as, (cf. e.g. Cohen [2])

$$\hat{X}_p = \left\{ \frac{K_p}{r} \sum_{i=1}^{n} x_i^{\hat{\beta}} \right\}^{1/\hat{\beta}} \quad \text{(A2-2)}$$

where $\hat{\beta}$, the ml estimate of β, is the solution of,

$$\frac{1}{\hat{\beta}} + \frac{1}{r}\sum_{i}^{r} \log x_i - \frac{\sum_i^n x_i^{\hat{\beta}} \log x_i}{\sum_1 x_i^{\beta}} = 0 \quad \text{(A2-3)}$$

H has been shown[3], that the quantity

$$u(r,n,p) \equiv \hat{\beta} \log \frac{\hat{x}_p}{x_p} \quad \text{(A2-4)}$$

follows a distribution that depends only on r, n and p and not on the parameters X_p and β of the population being sampled. Percentage points of this distribution must be determined by Monte Carlo sampling. Similarly, the quantity,

$$v(r,n) \equiv \hat{\beta}/\beta$$

has been shown[3], to be distributed independently of the Weibull population parameters.

To test the hypothesis Ho: $x_p = (x_p)_o$ against the alternative H_1: $x_p = (x_p)_1 \leq (x_p)_o$ one computes the ml estimates $\hat{x}_p$ and $\hat{\beta}$ from Eqs. (A2-2) and (A2-3), and accepts the null hypothesis if,

$$\hat{x}_p > (x_p)_o \exp(u_\alpha/\hat{\beta}) \quad \text{(A2-5)}$$

If Ho is true, i.e. if $x_p = (x_p)_o$, the probability of accepting Ho under this procedure is $1-\alpha$. When H_1 is true, the probability of acceptance is less than $1-\alpha$ by an amount that depends upon the quantity $\beta \log {(x_p)_o}/{(x_p)_1}$ (or equivalently on $[x_{p_o}/x_{p_1}]^\beta$

This follows from,

$$P_a = Prob\left[\hat{x}_p > (x_p)_o \exp(u_\alpha/\hat{\beta})\right] \quad \text{(A2-6)}$$

$$P_a = Prob\left[\hat{\beta} \log \frac{\hat{x}_p}{(x_p)_o}\frac{(x_p)_1}{(x_p)_1} > u_\alpha\right]$$

This becomes, after rearranging and using $u = \hat{\beta}\log \hat{x}_p/(x_p)_1$

$$P_a = Prob\left[\frac{u_\alpha - u}{v} < -\beta \log \frac{(x_p)_o}{(x_p)_1}\right] \quad \text{(A2-7)}$$

Defining $S(\alpha, r, n, p) \equiv \frac{u_\alpha - u}{v}$, it is observed that the distribution of s, being a function of u and v is distributed independently of the Weibull parameters.

Its percentiles may also be determined by Monte Carlo sampling.

For a fixed probability γ of accepting Hp, ($P_a = \gamma$) one has

$$\beta \log {(x_p)_o}/{(x_p)_1} = -S_\gamma \quad \text{(A2-8)}$$

where S_γ denotes the 100 γ-th percentage point of s(γ,n,r,p).

Alternatively one may express Eq. (A2-8) as,

$$\left[{(x_p)_o}/{(x_p)_1}\right]^\beta = e^{-S_\gamma} \quad \text{(A2-9)}$$

b) Confidence Interval Criterion

Two sided 100(1-α)% confidence limits for X_p follow from the probability statement,

$$Prob\left[u_{\alpha/2} < u = \hat{\beta}\log\frac{\hat{x}_p}{x_p} < u_{1-\alpha/2}\right] \quad \text{(A2-10)}$$

and are,

$$\hat{x}_p \cdot \exp(-u_{1-\alpha/2}/\hat{\beta}) < x_p < \hat{x}_p \cdot \exp\left(-\frac{u_{\alpha/2}}{\hat{\beta}}\right) \quad \text{(A2-11)}$$

The ratio of the upper to lower confidence limits is,

$$R = \exp\left[\frac{u_{1-\alpha/2} - u_{\alpha/2}}{\hat{\beta}}\right] \quad \text{(A2-12)}$$

R is seen to be a random variable since it involves the random variable $\hat{\beta}$. The median value $R_{0.50}$ of R, is computed by substituting $\hat{\beta}_{0.50} = \beta v_{0.50}$ to give,

$$R_{0.50} = \exp\left(\frac{u_{1-\alpha/2} - u_{\alpha/2}}{\beta v_{0.50}}\right) \quad \text{(A2-13)}$$

Alternatively, one may write,

$$\beta \log R_{0.50} = \frac{u_{1-\alpha/2} - u_{\alpha/2}}{V_{c,50}} \qquad \text{(A2-14)}$$

The quantity $\beta \log R_{0.50}$ is a function of r, n and p inasmuch as it is computed from percentage points of u(r, n, p) and v(r, n).

References

1. Epstein B., and Sobel, M., "Life Testing", JASA, 48, pp. 486-502, 1953.

2. Cohen, A. C., "Maximum Likelihood Estimation in the Weibull Distribution Based on Complete and on Censored Samples", Technometric, 7, pp. 579-598, 1965.

3. McCool, J. I., "Inference on Weibull Percentiles and Shape Parameter from Maximum Likelihood Estimates", IEEE Trans. on Reliab., R-19, pp. 2-9, 1970.

4. McCool, J. I., "Inference on Weibull Percentiles from Sudden Death Tests Using Maximum Likelihood", IEEE Trans. on Reliab., R-19, pp. 177-179, (1970).

Acknowledgement

Most of the work presented herein was sponsored by SKF Industries, Inc. Permission to publish is gratefully acknowledged. Some of the numerical values were developed under the sponsorship of the Aerospace Research Laboratories, Air Force Systems Command, United States Air Force, Wright-Patterson AFB, Ohio 45433, Contract F33615-72-C-1761.

TABLE 1

VALUES OF $\left[(x_{0.10})_0/(x_{0.10})_1\right]^{\beta}$ FOR ACCEPTANCE PROBABILITIES OF 10% AND 50%

Sample Size n	No. of Failures r	$u_{0.10}(r,n,0.10)$	$\left[(x_{0.10})_0/(x_{0.10})_1\right]^{\beta}$ For 10% Acceptance: Weibull	Exponential	$\left[(x_{0.10})_0/(x_{0.10})_1\right]^{\beta}$ For 50% Acceptance: Weibull	Exponential
10	3	-1.03	4.81	4.83	2.48	2.43
	4	-0.77	4.50	3.85	2.24	2.11
	5	-0.72	4.48	3.29	2.26	1.92
	7	-0.70	4.46	2.71	2.30	1.71
	10	-0.68	4.39	2.29	2.24	1.56
20	3	-1.60	8.47	4.83	2.82	2.43
	4	-0.86	3.82	3.85	2.11	2.11
	5	-0.69	3.37	3.29	1.93	1.92
	6	-0.61	3.12	2.94	1.86	1.80
	7	-0.59	2.64	2.71	1.85	1.71
	10	-0.54	2.92	2.29	1.90	1.56
	15	-0.54	2.89	1.96	1.67	1.42
	20	-0.52	3.07	1.78	1.82	1.35
30	3	-2.59	32.3	4.83	3.98	3.43
	4	-1.27	5.46	3.85	2.36	2.11
	5	-0.80	3.44	3.29	1.94	1.92
	6	-0.65	3.00	2.94	1.81	1.80
	7	-0.57	2.88	2.71	1.75	1.71
	8	-0.53	2.72	2.52	1.72	1.64
	9	-0.51	2.72	2.39	1.71	1.59
	10	-0.49	2.69	2.29	1.67	1.56
	15	-0.46	2.65	1.96	1.68	1.42
	30	-0.43	2.56	1.60	1.62	1.28

TABLE II

VALUES OF $R^{\beta}_{.50}$ FOR 80% CONFIDENCE LIMITS ON $X_{0.10}$ AND $X_{0.50}$

Sample Size n	No. of Failures r	Median 80% Confidence Limit Ratio, $R^{\beta}_{.50}$ Weibull Case $X_{0.10}$	$X_{0.50}$	80% Confidence Limit Ratio, R Exponential Case X_p
10	3	20.8	2.2	4.83
	5	13.4	4.2	3.29
	10	8.0	2.7	2.29
15	5	8.2	5.7	3.29
	10	6.2	2.4	2.29
	15	5.1	2.2	1.96
20	5	6.2	8.0	3.29
	10	5.2	2.5	2.29
	15	4.6	2.1	1.96
	20	4.1	2.0	1.78
30	5	4.4	12.8	3.29
	10	3.8	2.9	2.29
	15	3.6	2.1	1.96
	20	3.5	1.8	1.78
	30	3.1	1.7	1.28

FIG. 1
TENTH PERCENTILE RATIO VS. NO. OF FAILURES

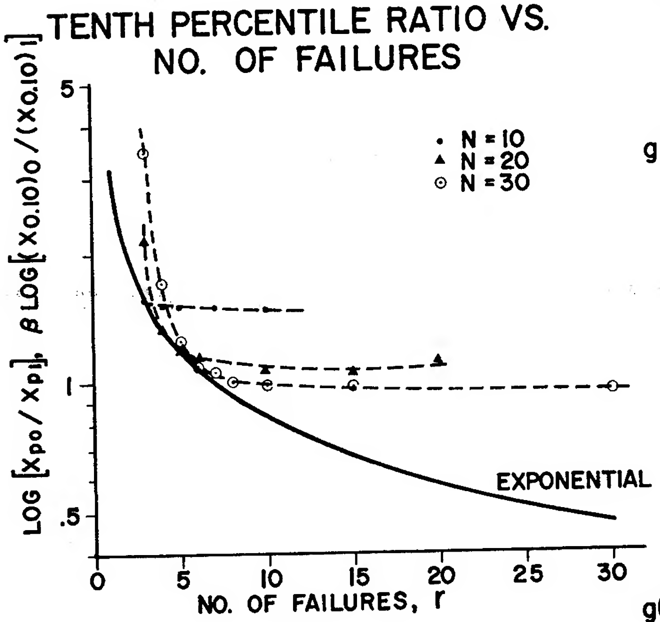


FIG. 2
FAILURE AND ESTIMATE DISTRIBUTIONS

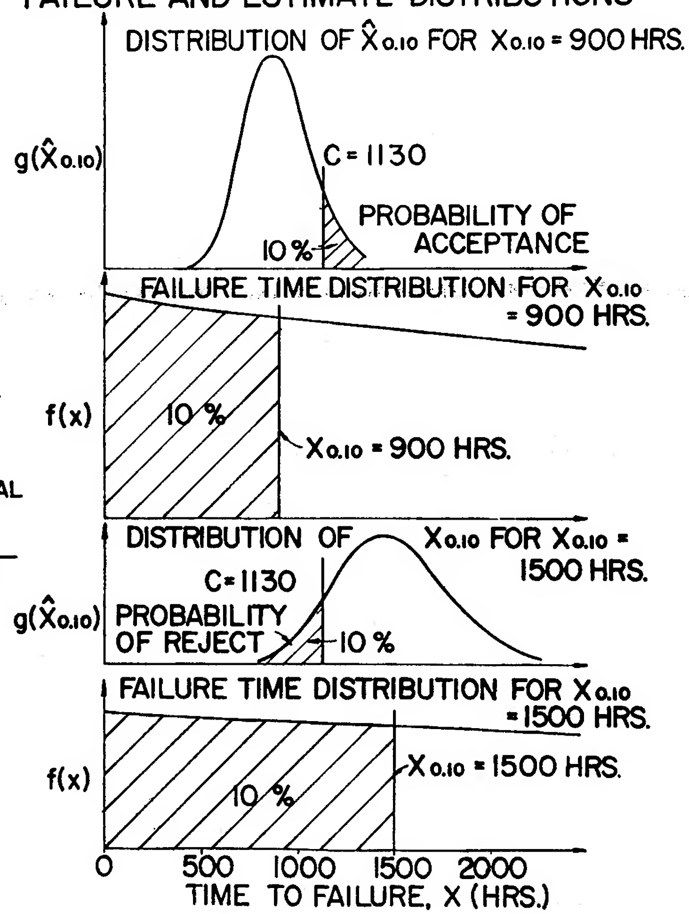


FIG. 3

MEDIAN CONFIDENCE LIMIT RATIO ($X_{0.10}$)

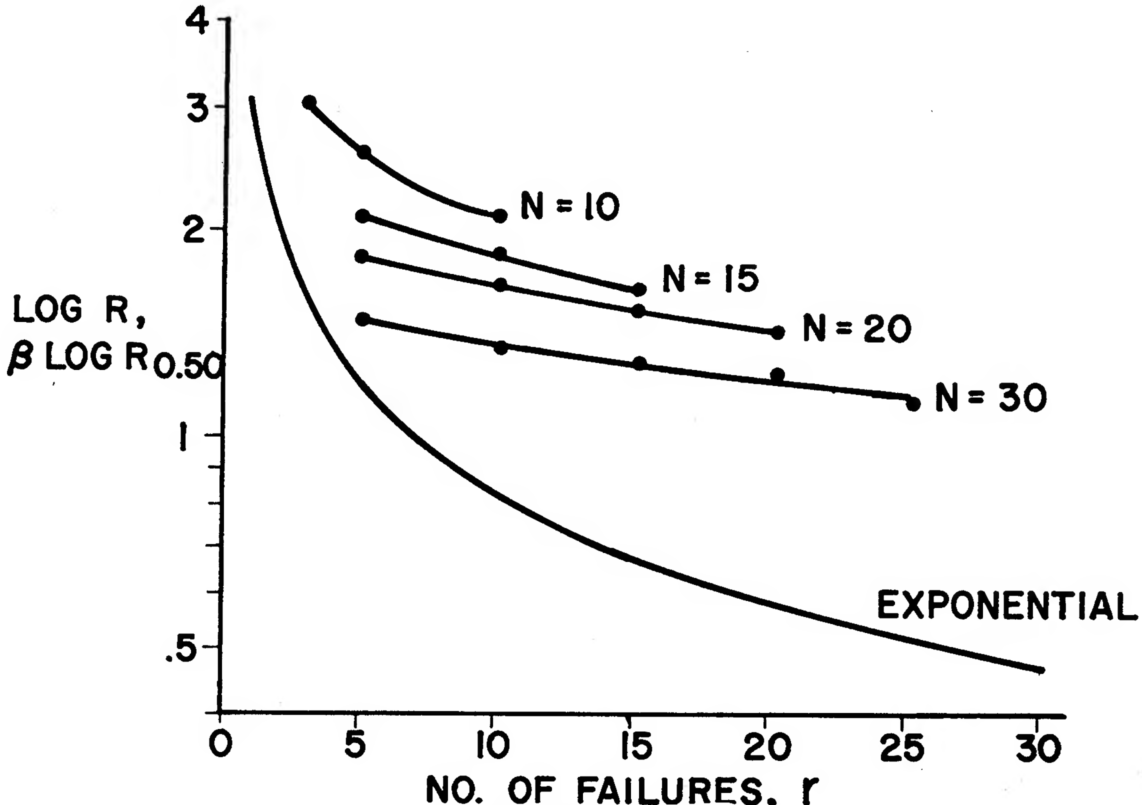
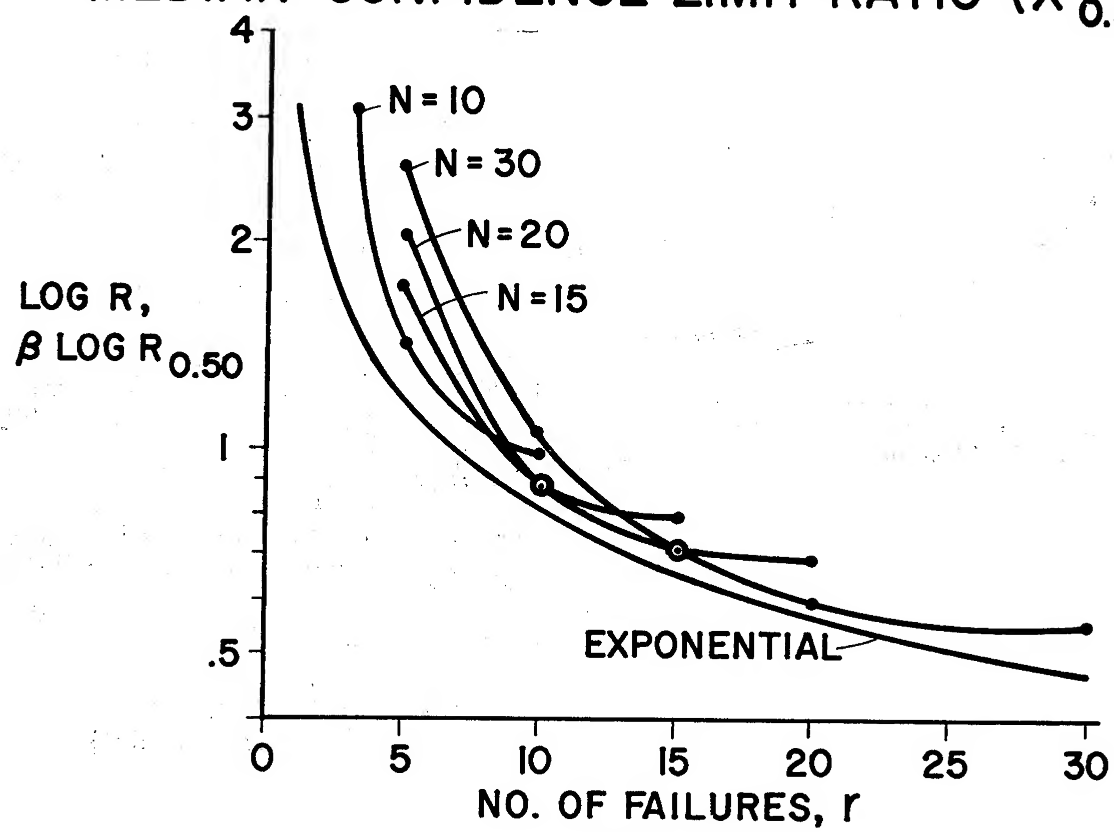


FIG. 4

MEDIAN CONFIDENCE LIMIT RATIO ($X_{0.50}$)



PANEL DISCUSSION OF BAYESIAN ANALYSIS IN RELIABILITY

Introductory Remarks

by Frank Proschan, Moderator
Florida State University

Bayesian methods have been proposed for and applied to a wide variety of reliability problems. In this panel discussion we hope to answer (at least partially) the following basic questions concerning the controversial Bayesian approach to reliability:

(1) What is the Bayesian approach? How is it used in reliability problems?

(2) Is it better than the classical approach?

(3) What are its weaknesses and strengths?

(4) Is the input data it requires available in actual reliability situations?

(5) Does the Bayesian approach lend itself to abuse? Does it invite the reliability analyst to assume information he does not really have?

(6) Has the Bayesian approach been successfully applied to real life reliability problems? What case histories can be described in which Bayesian methods were used with benefit?

(7) Does the Bayesian approach lead to a unified theory of statistics? If so, in what way is this of concrete value to the reliability analyst?

For the benefit of the few who have somehow miraculously escaped the Bayesian storm of controversy, we summarize the essence of the Bayesian approach in a reliability context. Assume system lifelength is governed by probability density $f(x|\theta)$, where θ is a population parameter of interest, such as mean lifelength. Under the classical approach, θ is unknown but fixed. Under the Bayesian approach, θ itself is a random variable with a priori density say, $g(\theta)$ before any observations are taken. Then the a posteriori probability density of θ, after an observation x on system lifelength is made, is given by

$$g(\theta|x) = \frac{f(x,\theta)g(\theta)}{\int f(x,\theta)g(\theta)d\theta}.$$

$g(\theta)$ may be interpreted as a measure of belief that system mean lifelength is θ before taking an observation, while $g(\theta|x)$ is the modified measure of belief that system lifelength is θ after taking the observation x.

The Bayesian asserts that prior information about θ may be quantitatively utilized under this formulation; prior information might come from previous experience with similar systems, engineering opinion, understanding of the physics of failure, a systems analysis based on component information, etc. The classical statistician objects that to quantify one's personal belief and incorporate it into a statistical analysis may lead to highly subjective answers that inextricably merge qualitative personal belief with quantitative scientific observation.

Perhaps a sensible approach for the reliability analyst to take is to use the Bayesian approach where prior information can be reasonably quantified and to use classical statistics otherwise. For example, in acceptance sampling of lots, a good deal of quantitative information is usually available concerning previous lots submitted; i.e., $g(\theta)$ may be safely estimated. On the other hand, if no quantitative a priori data is available, but a particular form of $g(\theta)$ is suggested only because it constitutes a natural conjugate prior, then the reliability analyst may well be suspicious of the resulting Bayesian analysis.

I stop at this point, since by now I have undoubtedly alienated both Bayesian and classical analysts.

BAYESIAN METHODS IN RELIABILITY

Frank Grubbs
Aberdeen Proving Ground

Many years ago, perhaps in the early 1950's, the late Professor Samuel S. Wilks told me that in the history of statistics Bayesian methods seem to run on a thirty-year cycle, but sometime before Sam's passing in 1964 we discussed the problem of the thirty-year itch again, and we were both convinced that the methods of Bayes were here to stay. Indeed, why not? At any rate, there is always the pressure for Ph.D. dissertation topics on anything that can possibly be dreamed up, and those interested in Bayes techniques could try and apply such principles to just about every problem which the so-called "classical" methods had already been applied to. (The author never understood why Bayesian techniques were not branded as "classical" originally on one hand, and what is now called "classical" should rather be referred to as "modern".) Oh my! With this little introduction, I may have already branded myself as "subjective" and not "objective", whereas I am a "frequentist". In any event, it is not surprising that by now we should have expected many, many applications of the principles of Bayes to various reliability problems, they have indeed occurred, and everyone including the followers of Bayes feel so much more comfortable when the Bayes procedures give the same answers as the exact classical methods! Does anyone have great confidence in the selection of prior distributions of parameters?

Perhaps some of the more important problems in reliability include various analyses of system reliability and the placing of confidence bounds thereon. In fact, since many tests of whole systems are too costly, or sometimes destructive, then it is advisable to estimate confidence bounds from component test data under laboratory conditions or inexpensive simulations of service conditions. Laboratory tests of increased severity and the like, if they can be related in some way to service conditions, may be used to predict system performance in the actual service environment. Also, tests on prototypes of systems, where such trials must be made, may be used to obtain data which can be employed to predict confidence bounds on system reliability in the intended or expected environment. Both classical and Bayesian methods may be used to "solve" many of these problems. Which approach should be used? Which can be depended upon and when is a satisfactory answer obtained?

We must remember that when one is clever enough to make the classical approach work for reliability problems, for example, even when it is necessary to transform the intricate statistical formulations around

into appropriate conditional probabilities, there generally can be no valid complaints, for otherwise who would accept the Bayes answer? (Don't accuse me of being a Bayes advocate when I select the classical approach "subjectively", as I know it is the right thing to do!) In such cases, there is no need for the Bayesian approach, except his proponents might well claim that their theory is "simpler" in some cases. In some reliability problems the classical methods do indeed get bogged down to some extent because of the complexity which creeps in, and I mention, for example, the problems of placing confidence bounds on complex system reliability for the case of components having binomial pass-fail or exponential time-to-fail data. On the other hand, the Bayesian approach requires considerable or ethereal intuition, or just plain ESP (!), to find the appropriate prior distribution which gives the correct answer, and how does one know he has the correct answer unless he checks it against the classical finding, which as we have said it is sometimes more difficult to find! Perhaps this latter observation helps the Bayesian cause in such cases. Something about the Bayesian approach which leaves me cold usually is that the improper prior in so many cases is the best to use! For example, for a series system and exponential time-to-fail data for components, the uniform assumption on the prior distribution of component parameter failure rate gives strange results for confidence bounds on overall system reliability, too low (high) for high (low) reliability! Nevertheless, the improper prior, or the reciprocal of component true failure rate, is the correct prior leading to optimum exact bounds for a single component system, and for $k > 1$ components in series, this Bayesian assumption leads to equivalent findings from Monte Carlo simulation methods which generate component failure rates which follow Chi-Square distributions. But even here system reliability bounds are obtained which turn out to be equivalent to the simple fiducial approach of Grubbs (1971). Why bother with Bayesian techniques in this case therefore? Which brings me to the next important point. Now please do not say that there are enormous amounts of data on file from which the priors can "easily be established". There are indeed much, much data on file, but such were not taken for the purpose of predicting priors for the Bayesians to use, the data are rather spotty in fact and not easily analyzed for such purposes either. Also, when the Bayesian advocates do not use simple conjugate priors, their work becomes rather intractable. This should be enough to raise grave doubts about Bayesian methods in reliability.

The writer's experience in solving reliability problems in an acceptably practical manner is that neither the classical nor the Bayesian methods necessarily work too easily or well. In fact, the fiducial approach (which certainly is not Bayesian) has much to offer, as it is more classical and one has simply to pivot on the sample data and then see just how far population values could wander away. Now you can see that if only the classical and the Bayesian people fight each other, then such two-sided conflict makes life really simple and easy. But how would you like to be an advocate of the fiducial approach so that both the Bayesians and the classicists take great delight in jumping on you together? Hence, the problem is such that one cannot talk only of Bayesian methods, but it must be enlarged to include the whole gamut of methods of solving statistical type problems, especially in reliability. Sometimes the same answers are obtained from the classical theory and the fiducial theory, as for the two-parameter negative exponential distribution and reliability confidence bounds thereon, but in other cases, such as the Behrens-Fisher problem, the fiducial method falls down. On the other hand, certain prior distributions for Bayesian approaches do indeed give the same posterior distribution as the fiducial approach, so that there is some connection there too. A case in point is that of exponential series system reliability for a fixed number of failures per component mentioned above. But who would prefer either the Bayes or the fiducial approach above the classical method? Not me!

Perhaps, we are merely saying that it pays for intelligent people, or even intelligent adversaries, to communicate if progress is to be made, and in the end we had just as well use those techniques which work well in practice, or check each other out, and are sufficiently simple and understandable - a very desirable goal, is it not?

In summary, the "Operations Research" approach may be needed for the problems of Bayesian Methods in Reliability!

THE ROLE OF BAYESIAN METHODS IN RELIABILITY

R. E. Schafer
Hughes Aircraft Company
Fullerton, California

First, to provide a framework for discussion, we will briefly contrast the Bayesian and Classical methods. This is probably best done by noting the input variables required by Bayesian methods that are not required by Classical methods: i) the prior distribution and ii) the loss function. Thus Bayesian methods permit decisions to be made on expected loss (as against the Classical method which uses confidence statements and (inductive) probabilities; in fact the foundation of the Classical method is the likelihood function) which is presumably the penultimate criterion in decision making. Subjectively and personalistically established prior distributions will be rejected out of hand for this discussion since, at this time, they do not provide a suitable means of communicating scientific and engineering knowledge. In short the prior probability distribution is to be interpreted in the usual frequency sense. However, we will call any method which uses a prior distribution a Bayesian method even though the decision criterion may not be based on a loss function.

Any statistical decision model, be it Bayesian or Classical, must rise or fall, on: i) how well it models the real world, ii) how convenient the model is to use, and iii) the cost efficiency of the method. Classical models have, to me, the sometimes misleading advantage that they are extremely convenient to use: one usually, in reliability applications, need only select a test size (significance level) or confidence level to make decisions. Consider as an example the popular MIL-STD 781B test plans. They are used with abandon: the requirement of a (conditional) exponential time-to-failure distribution being rarely if ever checked. These tests are not particularly robust under some alternatives. On the other hand, Bayesian methods model the real world well in two respects: if the parameter in question is a random variable this can be modelled by using a prior distribution and the decision criteria can be related to real life, e.g., loss functions or posterior probability rather than the usual "inductive" measures. In fact in "selling" Bayesian methods in reliability one often hears the advantage that prior knowledge can be incorporated into the decision process. This is true enough but it is interesting to note that the impetus to Bayesian models has always been, and continues to be, the attractiveness of the decision criteria, i.e., minimization of expected loss or posterior probability.

The Bayesian model clearly has useful applications in reliability practice for in many situations the parameter of interest may be considered a random variable. For example, a succession of "identical" computers clearly have different, although unknown, mean-times-to-failure. The problem is in establishing the prior distribution. Methods of fitting prior distributions to observed reliability data have been discussed[1]. If enough data to fit a prior distribution is not available, empirical Bayes methods can be used[2]. Regarding the loss function, it is sometimes difficult, if not impossible, to establish such a function to everyone's satisfaction. However, in reliability work, the posterior probability of the hypothesis $P(H|Data)$ has been commonly used instead of the loss function. This clearly seems better than the Classical $P(Data|H)$ which is inductive.

In view of the fact that Bayesian methods often model the real world better (than the Classical Models) the real growth of Bayesian models in reliability will be due to a reduction in decision costs. When Classical models are used in a "random parameter" environment costs of decision are often increased because one pays for protection one does not need. An example will serve to illustrate this last point. Suppose the parameter θ in an exponential time-to-failure distribution is a random variable and that a MIL STD 781B test is used with (minimum acceptable mean-time-to-failure) θ_1 and consumer's risk $\beta = .01$. Suppose now that its known that, a priori (i.e. before the test is taken) $P(\theta \leq \theta_1) = .001$. Then surely the event $\theta \leq \theta_1$ is so rare that it is not worth (speaking Classically) "paying" for the sample size to provide a β as small as .01. One is paying for protection he does not need and a $\beta = .10$ would surely have been acceptable.

References

[1]Schafer, R. E. and Feduccia, A., "Prior Distributions Fitted to Observed Reliability Data", IEEE Transactions on Reliability, August, 1972.

[2]Robbins, H., "An Empirical Bayes Approach to Statistics", Proc. Third Berkeley Symposium, Math. Statist. and Prob., University of California Press.

BAYESIAN STATISTICS

Nozer D. Singpurwalla
The George Washington University

The position taken by this panelist is akin to that of an attorney charged with the task of defending his client. As such, the approach will be one that rebutts the criticisms leveled against the so-called "Bayesians". Because of the position taken, it is requested that this panelist not be castigated in the future either as a "Bayesian" or as a "non-Classicist".

A review of the literature reveals that some of the arguments directed towards a reliability analyst (engineer, statistician) using Bayesian methods are along the lines indicated below. The position of this panelist on each of these issues is given after the issues have been presented.

1. The Bayesian approach lends itself too readily to misuse.

Such criticism is unfortunate because it tends to shadow the real issue "are Bayesian methods useful in reliability?" This panelist concurs with the point of view that merely the use of Bayes theorem does not make one a Bayesian. He also concurs with the point of view that those situations (such as acceptance sampling, life testing from distinct lots, etc.), which call for the use of a prior distribution on an unknown parameter, should not be classified as legitimate Bayesian procedures. There may be several other genuine misuses of Bayesian procedures as there are misuses of any other procedure, and such criticism is not considered serious enough for a strong rebuttal.

2. The assumption of a completely specified prior distribution is not rigid.

The classical argument leveled against a Bayesian is in his choice of a prior distribution. The position of this panelist is that all statistical procedures when applied to practical problems involve some element of personal choice. For example, in the testing of statistical hypothesis, the choice of type I and type II errors is left to the decision maker. In acceptance sampling the risks are negotiated between the producer and the consumer. There is no reason why the existence of a prior distribution cannot be included in a system of axioms, and the choice of this prior be left to a decision maker, or be negotiated between the parties involved. Any discussions and disagreements on the choice of suitable priors should not be confused with discussions on the legitimacy of Bayesian procedures.

3. Bayesian methods are of mathematical interest only and not applicable to the real-world problems of data collection and analysis.

This point of view is often expressed by those statisticians interested in "the return of statistics from the mathematician to the statistician". They adhere to the point of view that their sole function is learning from the data alone.

The above point of view is very short sighted, and perhaps disturbing even to our "non-Bayesian" panelists and our moderator. The strength of the statistical method lies in the fact that its procedures have been backed up by rigorous mathematical arguments. If all that reliability analysts did was collect and plot failure data and pass it up the line, then this panelist could be partially sympathetic to the argument. This not being the case, the above point of view needs to be challenged.

If coherent procedures of inference, based on any system of axioms, one of which could be the existence of a prior could be developed, then there is no reason why Bayesian procedures should be criticized. As a matter of fact, Lindley shows that any system of axioms which does not have a prior leads to procedures which are not coherent. Since the principle of coherence has a special significance in reliability theory, this aspect of non-Bayesian inference should not be overlooked.

ON THE USE OF BAYESIAN METHODS IN RELIABILITY

INDEX SERIAL NUMBER – 1081

Robert G. Easterling
Sandia Laboratories
Albuquerque, New Mexico

Workers in reliability are often faced with analyzing a body of data. The questions they ask themselves and the statistician often amount to, "What can I learn from these data?" Rarely, at least in my experience, do they ask, "What decision should I make?" or "How sure should I be that a parameter θ is between two specified numbers?" Yet these latter questions appear to be the ones addressed by proponents of the Bayesian approach. It is my position that not only are these artificial questions, but the Bayesian answers are endowed with a false sense of precision. Furthermore, I object to the Bayesian approach because it ignores fundamental statistical concepts such as the assessment of sampling variation and because it merges the information contained in the data with the data collector's opinions. It should be made clear here that by the Bayesian approach I am referring to the incorporation of personal probability into the analysis of data. I have no Bayesian quarrel in the situation in which a parameter, such as the MTBF of a lot, is a physical random variable. However, I think it is rare that one can completely specify the distribution of an unobserved random variable.

The way one learns from data, I believe, is by asking, "What sort of (probability) model is consistent with the data?" This question is the very essence of scientific investigation. The statistician seeks out the answer by considering the rules by which the data were collected, by making goodness-of-fit tests, and, if a parametric model can be fitted, by making statements about what parameter values are consonant with the data. Through experience, I have found this to be an informative, even enlightening, process. To introduce arbitrary elements, such as priors and loss functions, into the process and re-express it as a series of prior and posterior betting odds is a gross distortion.

The Bayesian approach requires that one express his prior beliefs about a parameter in terms of a completely specified probability distribution function. However, it seems to me that the best one can do is to specify a range or a distribution of his personal probability. That is, a personal probability is a random variable in the classical sense; it varies from day to day, moment to moment. To amplify, if you claim 19:1 prior odds on some proposition, I assert that it would be difficult to say why the prior odds couldn't be 18:1 or 20:1. It seems imperative that this source of variation be accounted for, but I have seen no evidence that Bayesians recognize it, much less try to account for it. This is why I claim a false sense of precision in the Bayesian approach.

It is often thought that the use of a "flat" prior represents vague prior belief. These priors can lead to results algebraically equal to results derived under classical approaches. However, this coincidence should not blur the distinctions. The classical confidence interval, in spite of its name, is just a statement about what range of parameters are consonant with the observed data to a specified extent. The "confidence" one has that the parameter is in that interval is confidence only in a very limited sense. The Bayesian posterior probability interval, on the other hand, throbs with confidence. It measures strength of belief and provides betting odds. The fact that this major change in interpretation and meaning is accomplished by only making the additional assumption of a flat prior suggests that this assumption is not at all vague, innocuous, or to be lightly taken.

Aside from my hesitancy to accept the assumptions necessary for implementation of the Bayesian approach, I wonder how I should interpret results expressed in terms of your personal probability. It appears that the only way I can do this is to search out your data. My understanding would be aided if you would summarize the data in ways such as goodness-of-fit tests and consonance intervals, which do not quantitatively incorporate your prior beliefs. The language of personal probability may be appropriate for describing the formation of personal opinion (talking to oneself), but it is not appropriate for the communication of scientific information.

The output of a Bayesian analysis is a statistic, a function of the data, and as such has operating characteristics. We learn from data by seeking to represent them as "typical" data generated by a particular model or class of models. The value, or use, of a statistic can be assessed by considering its operating characteristics, that is, its behavior in repetitions from the model. The Bayesian apparently sees no value in this. His posterior distribution or decision is his personal statement based on the data at hand and he cares not how that statement could vary with hypothetical data generated by the model he has chosen to represent his observed data. The assessment of sampling variation is widely regarded to be a cornerstone of data analysis and I cannot see why the incorporation of personal probability into the analysis should permit one to discard this basic concern.

To summarize, the Bayesian approach is based on often untenable assumptions, leads to non-communicable results, and distorts the process by which one learns from data. I see no reason to regard it as a substitute for or as a practical extension of classical statistical methods.

INDEX SERIAL NUMBER — 1082

OPERATIONAL READINESS AND MAINTENANCE TESTING OF THE B-1 STRATEGIC BOMBER

Jack E. Ogden
Supervisor, CITS Group
North American Rockwell Corp.
Los Angeles, California

Frank P. Cavanaugh
Lead Engineer, CITS Group
North American Rockwell Corp.
Los Angeles, California

Studies involving the operational use and role of the Advanced Manned Strategic Aircraft (AMSA) provided a set of requirements for inflight and ground testing which could not be met by conventional Aerospace Ground Equipment (AGE). With the evolution of the AMSA into the B-1 bomber, it became necessary to develop new concepts in on-aircraft testing. The implementation of these concepts led to the development of the Central Integrated Test Subsystem (CITS).

The CITS is an onboard aircraft subsystem utilizing a dedicated computer to read and assess the health of all airborne subsystems (including itself). This provides the aircrew with continuous status of the subsystems and the ability to isolate failures to a line replaceable unit (LRU).

One of the design goals for the CITS is to minimize the requirements for flight line AGE and to impact, where possible, the shop AGE.

This paper describes the CITS in its relationship to the B-1 and to its maintainability, with particular emphasis on cost trade-offs between the CITS and all levels of AGE.

Background

More than a decade ago, a need to modernize our bomber force was recognized. This resulted in a series of studies for various aspects of an AMSA being conducted by several aircraft, avionic, and engine companies. In turn, these studies culminated in a set of requirements for a new, versatile type of bomber which was designated the B-1. The AMSA studies clearly demonstrated that the B-1 would have to meet very different requirements than any previously designed bomber, such as the ability to fly at close to the speed of sound at very low altitudes under radar detection, reduced radar cross section, reduced infrared reflection, and the ability to use a much shorter runway than conventional bombers in order to be dispersed to many different types of airfields. This requirement made the conventional use of AGE impractical from the viewpoints of the quantities necessary to preposition AGE at all possible airfields; the cost of maintaining the larger inventory of AGE in terms of spares, manpower, time involved in getting to and from the equipment etc; and from the logistical view of needing to know what equipment is at what base and whether this particular AGE is configured for the next aircraft configuration it will be required to test. These and other considerations dictated a positive need for an onboard test system. Studies of existing (and contemplated) test systems in use by both military and commercial aircraft in that time frame indicated that the B-1 onboard test system would have to advance the state of the test system art in terms of reduced costs and weight and increased flexibility and effectiveness if it were to do the necessary job.

All of these factors combined to culminate in the test system presently under development by North American Rockwell which is known as the CITS, i.e., Central Integrated Test Subsystem. This acronym was selected to describe the system which is central in the sense of bringing data from all aircraft subsystems, and integrated in the sense of using sensors/transducers already installed in each of the subsystems for operational purposes, as the source of most of the CITS signals. A block diagram of this system is shown in Figure I.

Description

The CITS is defined as an airborne subsystem which keeps track of the operability status of all other airborne subsystems to provide flight and ground crews with timely and accurate information regarding subsystems "health." It comprises a central digital on-board computer which receives information from several Data Acquisition Units (DAU) and provides readouts in the form of lighted messages on the CITS Control and Display (CCD) Panel, printed information on an onboard printer, and at some later date, digital recording on magnetic tape. The concept of a CITS type system offers many advantages to the user such as (1) immediate failure indication; (2) increased flight safety, since the condition of the aircraft is known during all phases of flight; (3) increased mission effectiveness, for the same reason; (4) increased aircraft availability, since its status is known before, during, and after each flight; (5) reduced test time, which leads to increased equipment life, since it is being tested while it is normally operating or for much shorter periods of time on the ground; (6) faster repair time, which leads to reduced maintenance man-hours per flight hour (MMH/FH); (7) maintenance as required, either through failure indication or trend analysis, which will allow more flight time; (8) reduction in incorrect fault diagnosis, which can lead to a more favorable logistics posture; (9) the high degree of automaticity which will assist in reducing the requirements for highly specialized personnel training.

Physically, the CITS consists of an onboard digital computer receiving inputs from a series of DAU's and providing outputs to a CCD panel, a clear text printer, and a magnetic digital tape recorder.

The CITS computer is the heart of the testing system since it provides the control mechanism to

collect test parameter data and to route information to the panel, the recorder, and the printer. It provides the "intelligence" which runs subsystem tests by bringing signals out of the aircraft subsystem at the required time and rate and performing the necessary evaluations to determine the status of each subsystem in its inflight relationship to all other interfacing systems. Briefly, the computer has a repertoire of 70 instructions, is capable of performing 200,000 logical operations per second, has a memory capacity of 32,768 sixteen-bit words, and can perform 1,000,000 memory operations per second. Figure II shows the computer and some additional data related to it.

This test system's interface with the outside world is principally through the CCD Panel which is located in the aft crew station for the use of the Offensive and Defensive Officers during flight, although the forward station crew members are provided with the capability to initiate inflight fault isolation tests. The CCD will simultaneously indicate status of the air vehicle subsystems and/or modes of operation of these subsystems while in flight and during ground readiness tests. One or more legends will be illuminated whenever one or more of the air vehicle subsystems go outside of their performance limits. An absence of illumination will indicate a GO system, and lamp testing will be performed to be certain that all of the redundant lamps have not failed. A ten-digit pushbutton input panel, which is a part of the CCD, is used to initiate all fault isolation testing when a manual request is to be made, in addition to providing a centralized location for the CITS man-machine interface. The approximate dimension of the panel and a view of the display portion are shown in Figure III.

The CITS airborne printer is the principal tool of the ground maintenance crew in that it provides a hard copy printout of all CITS derived data. All fault detection data, i.e., subsystem failures, LRU identification and functional failure, in addition to fault isolation data, such as fault isolation instructions and isolation identification numbers, are printed on the tape output of this printer, thus providing the ground maintenance men with a complete picture of what occurred during the flight just completed. The printer will operate up to 25 characters per second on a 100-foot roll of tape.

In order to capture data which occur only during flight at a particular speed and altitude, a magnetic-tape, digital data recorder will be provided as a part of the CITS. The tape will be a cartridge replaceable module easily accessed from the ground using the crew entry ladder.

The data collected by the recorder will be available for use in Air Force Ground Data Processing Systems of the future. Figure IV shows a view of the printer and recorder in relation to their ease of access.

These three devices - the Display Panel, the Printer, and the Recorder - provide CITS with the capability to interface with the peculiar operational and maintenance environment dictated by the mission prescribed for the B-1 bomber. An outline view of the B-1 is shown in Figure V with the approximate location of the various components of the CITS.

The requirements of this mission demanded that the design of the CITS be different than existing test systems, in that design emphasis was required in areas usually (in the past) not considered as being very important. The design of the CITS had to be such that its operation was basically automatic and very simple to operate, with no handbooks required and no coded entries or number sequences to be deciphered before knowledge of a fault is gained. It was also necessary to minimize the use of Built-In Test Equipment (BITE) in spite of the fact that the aerospace industries and the military have been (and are) actively developing BITE. This was necessary since a hardware change in a subsystem could cause a change in the BITE and possibly a change in the rest of the test system, thereby requiring extensive hardware changes to the aircraft with a corresponding loss of utilization of the aircraft. The extensive use of end-to-end dynamic testing was one of the major design goals of the CITS in order to provide maximum generation of information with the smallest number of test points.

The main thrust of the CITS design, however, was flexibility - flexibility to allow software to control the testing functions rather than hardware since it is more cost effective to make a software change than to effect a hardware change to follow the air vehicle subsystem changes - flexibility in providing random access to test point data in order to allow the computer to "see" the data it needs when it needs it and not be forced to wait until that particular piece of information is available - flexibility in that the design of the CITS is not customized to fit one specific subsystem, but instead could be used in almost any type or configuration of a subsystem - and flexibility in its capability of growth to allow for additional subsystems to be added to the B-1 without requiring a new CITS system. These design goals have been and are being implemented in the mechanization of the CITS in order that this advanced test system become a major factor in lowering the MMH/FH ratio, in raising the number of hours of availability for each CITS-equipped aircraft, and in providing operational information to increase the probability of mission success.

CITS Usefulness

The different and somewhat unusual requirements of the B-1 bomber should be sufficient justification for developing a CITS type system but, in order to be completely effective, the CITS must offer a plus value for the maintainability of the aircraft or the personnel charged with the responsibility for maintaining the B-1 fleet will slowly, but inevitably, discontinue its use and will find other methods of testing the aircraft. To determine whether or not this plus value existed, it was necessary to study the disadvantages and the advantages inherent in the design and use of the CITS to be certain that the disadvantages did not outweigh the advantages.

Disadvantages

Since the CITS is an onboard test subsystem as compared to carry-on or roll-up equipment, its disadvantages were the same as all other subsystems internal to the aircraft. In this instance, the usual factors of weight, volume, power consumption, and cooling, and the effects on reliability and maintainability were considered.

The weight disadvantages was minimized through the extensive use of existing signals generated by the normal operation of the aircraft's subsystem rather than using special transducers whenever a parameter was required. Additional weight savings resulted from the multiple use of each parameter (where possible) instead of adding sensors/transducers or circuitry to provide additional parameter data. Fewer transducers and circuits resulted in less wire and connectors to further hold down the weight. Additional weight decrease was realized from the reduction in use of BITE, which could be removed from existing subsystems since it was no longer needed. Volumetric requirements were kept to a minimum through the use of aircraft components and equipment, most of which are designed for lightweight, low-volume use, and, as is the case for weight, through the multiple use of transducers, thereby reducing the need for add-on transducers. Power consumption and cooling requirements were also minimized through the use of these types of components, i.e., low-weight low-volume aircraft items. The use of these components in the design of the CITS hardware resulted in a test system which will weigh approximately 165 pounds, and use about 1100 watts of power and 4 cubic feet of space.

The overall reliability of the aircraft was reduced because of the additional circuitry and components added to the aircraft. However, this reduction was within the reliability requirements specified for the B-1. The penalty due to the reduction in reliability was offset considerably by the increase in probability of mission success due to the ability of the CITS to detect failures in flight thus providing the air crew with the knowledge needed to evaluate their position relative to their aircraft and its ability to continue the mission.

Maintainability of the B-1 was also impacted by the addition of CITS since it was another onboard system to be considered in the maintenance analysis. However, since CITS is basically a maintenance system and its contribution to the overall maintainability of the aircraft is so great, it is difficult to evaluate this system on the same basis as the other onboard aircraft subsystems.

Advantages

It would seem that the operational requirements which had to be fulfilled with some type of test equipment different than conventional AGE would be sufficient reason for the design and development of the CITS. However, maintenance considerations have to be looked at in order to determine that a CITS type system will provide advantages over Detached Test Equipment since it is possible that the operational requirements may change or be overshadowed by cost problems thereby forcing a decision between the operations and maintenance functions. Therefore, CITS will be described in relationship to AGE at all levels, and in its potential impact on reliability, maintainability, mission readiness (completion), and on the cost of ownership.

Advantages/AGE Impact

Organization AGE

A design goal of the CITS is to minimize (or eliminate) organizational or flight line AGE. Practical implementation of CITS indicates that the CITS should reduce the flight line test equipment about 85% of what would normally be used on this type of aircraft. While total elimination of this level of AGE is theoretically possible, practical experience strongly suggests that it cannot be done. However, during the entire RDT&E program, this goal will continue to be looked at in order to bring the above percentage figure down as low as constraints permit. As every AGE design engineer knows, the task of providing AGE at the same time that flight testing begins is one of the major problems confronting him and he is usually forced to use factory test equipment or some type of Special Test Equipment to bridge the gap between first flight and first need for his organizational type AGE. Even though CITS will be a development system at this point in time, previous shop testing for compatibility and integration will contribute a certain degree of confidence to its usage, and will allow it to be used as a means of testing the air vehicle subsystems. In addition to its early availability, the CITS will also have inherent advantages over conventional organizational AGE, in that the on-board equipment will be looking at actual operational signals produced in the environmental atmosphere of the subsystems. These same environmental characteristics are either simulated by flight line AGE or are ignored, thereby reducing the credibility of the test results. In many instances, it is impossible to exactly duplicate the unique combination of speed, altitude, temperature, onboard generator voltage variations, acceleration, hydraulic systems variations, and aircraft stresses which will produce the failure found in flight.

This capability of onboard testing will lead to fewer false failure removals and to a positive assurance that a subsystem statused as good is "good". This advantage will in turn provide an increased availability of the aircraft, permitting a higher utilization rate and lowering the MMH/FH ratio.

Field/Shop AGE

The impact of the CITS on the field/shop level of AGE will of course not be as great as on the organizational level since the shops normally are concerned with Shop Reparable Units instead of LRUs, and the CITS is LRU oriented. The impact should be felt,

however, in the areas of a need for less equipment since each LRU removed from the aircraft will be a defective unit which will lead to a reduced demand for shop test equipment since little or no machine time will be wasted checking out good units and recertifying them after test. This could also lead, less directly however, to a decrease in the amount of spare parts which have to be stocked; a decrease in the number of personnel required to operate a base shop, i.e., less time wasted in setting up and dismantling the units to be tested and less handling-shipping-locating-recording time, etc; and an increase in the availability of each of the LRUs. When the field shop and the aircraft are located at the same air base, a considerable amount of time and test equipment could be saved by performing the recertification of each failed and repaired LRU on the aircraft using the CITS as the testing source. This method of recertification would reduce the shop time required to test and retest a given LRU which indicates GO on the test set and NO-GO when inserted into its using system on the aircraft.

Depot AGE

The depot will be impacted the least by CITS since it is furthest removed from the operational aircraft and is usually not very sensitive to individual aircraft failures. However, there is still the secondary benefit of reducing the number of units in the "pipeline" since all (at least theoretically) of the units sent to the depot will be failed units and not good units marked as bad. This should reduce the time required to turn equipment around to return to the base because of a greater availability of the depot test equipment and due to fewer units in the pipeline. Most of the present-day AMAs are highly automated repair facilities and are becoming more so, resulting in a requirement for a high volume of failed equipment to be cost effective. This requirement has led to the consolidation of several support functions into a single AMA instead of their being spread over several different depot, thereby, in some instances, causing the failed units to be transported long distances to and from the using site. Therefore, since the CITS will provide a high degree of assurance that every LRU shipped to the depot has indeed failed, the depot equipment will be more cost effective by virtue of a high repair rate in conjunction with its utilization rate with a minimum of lost time trying to find failures in good units. This cost effectiveness must also consider such factors as unnecessary transportation costs for good units thought to be bad, time required for personnel connecting and disconnecting the unit under test, time required to pack and unpack the units, and time involved in record keeping during all phases of a unit's passage through the depot.

Advantages/Maintainability

To properly implement the CITS concept, it was necessary to prepare a list of LRUs for every air vehicle subsystem. This seemingly simple task became one of the most difficult problems in CITS implementation, since it was almost impossible to provide a general definition for an LRU for the non-avionic subsystems - the avionic subsystems should be a little less difficult because of the usual packaging concept used in avionic equipment. After several unsuccessful attmpts to define an LRU, it became obvious that each subsystem had to be looked at individually with complete regard for the way it was packaged and with concern for the maintainability requirements of the subsystem. As a note at this point, it should be mentioned that the ideal situation would be to give the onboard test system engineer the option to lay out the packaging of each air vehicle subsystem - this is something to dream about but probably never to be realized. From the above, it is apparent that the CITS impact on maintainability was not as great as it might have been on the physical access portion of the "ility", but - less obviously - CITS will have a major impact on other aspects of maintainability; i.e., reduced number of man-hours required to isolate faults in the aircraft's subsystems and to requalify them after repair; less test equipment required since CITS makes maximum use of test points by using data from one test point in several tests; fewer and less skilled maintenance men required; no time required for test equipment set-up and disassembly; less time required in actual checkout of the airborne systems since CITS is automatic and immediately available; reduced downtime of the aircraft providing a greater rate of utilization; fewer delays and aborts; and higher sortie completion rate.

All of the above factors contribute favorably to the most important of the B-1 factors and that is mission readiness, which is easily converted to mission completion. Any aircraft which has the capability of running a complete check of all of its non-avionic (and at a later date all of its avionics) subsystems as it taxis toward takeoff, will have a much greater probability of starting and completing its assigned mission than an aircraft which has not had a complete checkout since its last periodic inspection. This rapid checkout - less than a minute for subsystem level detection - will contribute immeasurably to the crew confidence that their upcoming mission will be successful, which in itself will generate an atmosphere of confidence with a corresponding increase in the probability of success.

Advantages/Cost Impact

While the probability of mission success is undoubtedly the single most important parameter, it becomes almost an academic point if the weapon system is too costly to operate or maintain. The CITS impact on the cost of maintaining the B-1 was studied closely during the AMSA period to be as certain as possible that the CITS offered a monetary advantage in addition to the other advantages as discussed above.

The cost study identified and evaluated the effect of CITS on factors such as personnel skill requirements, training, AGE spares requirements, numbers of maintenance men required, test equipment setup time required, and other similar elements. These factors were then converted into dollars in order to determine a ten-year cost picture which was

used to measure the impact on the CITS. As one part of the evaluation, it was necessary to make a determination of the test equipment which would be required to support an aircraft which does not have an on-board test system such as the CITS, as well as determining that test equipment which is required to support the B-1 in addition to the CITS.

The selected factors used to conduct this study and a brief description of each is presented here to indicate the depth of analysis used. Production start-up costs were considered to be non-recurring costs resulting from beginning production on a new system. Since CITS is a single system configuration, these start-up costs were lower by almost 80% compared to the many different types of AGE which would be required. Production costs, as opposed to production startup costs, were analyzed for one set of plane-side AGE and compared to one CITS system in production quantities. The AGE production costs include computations to equalize the multiple use of AGE (in this instance 3 aircraft per one set of AGE was used as the factor) and the single use of CITS (1 set per aircraft). It appeared that the CITS would cost more by about 7%. The AGE costs incurred during the Research , Development, Test, and Evaluation phases of the program were based on XB-70 data updated and adjusted to represent today's advanced technology. The CITS costs were based on existing similar test systems, increased to recognize the advancement to the state-of-the-art necessary to develop CITS. The AGE costs, because of the many and diverse types required, were higher than the CITS costs by over 33%. The number of maintenance men required and the cost of a maintenance man-hour, multiplied together, provided a factor used in computation of the average monthly cost of maintaining the aircraft. Since CITS impacts the skill level required by the maintenance man, this factor and the cost of one airman of this skill level were also determined for use in computing the cost of maintaining the B-1. Test equipment setup time, which is the time required to connect and disconnect the test equipment to the system to be tested and to connect and disconnect the ground supply equipment such as electrical, hydraulic, pneumatic, and cooling, is included in the cost study as a factor in the cost per time in maintaining the aircraft. The setup time was estimated to be 1 hour for AGE compared to no time for CITS.

The cost of test equipment spares was compared, since this will be a large factor in either type of test system. The cost of CITS spares appears to be approximately 10% higher than that of AGE probably due to the fact that CITS will require airborne certified spares and the AGE will not.

One major factor was the cost of additional AGE required if the CITS were not developed for the B-1. This cost did not include any of the other AGE costs used in other parts of the analysis, just those costs which could be clearly identified as being needed for replacement of testing functions performed by the CITS.

All of the factors described above were used in a somewhat simple straightforward equation to arrive at a dollar value for two separate B-1 configurations, one equipped with the CITS and the other equipped with a small amount of BITE, but mostly relying on conventional AGE for checkout and maintenance. The results of this study show that the B-1 equipped with CITS will provide a cost advantage - in addition to the other advantages discussed above - over a B-1 utilizing conventional AGE and will, over an estimation period of ten years, lower the cost of ownership approximately 12.5% which in most military aircraft systems would be measured in the hundreds of millions of dollars. While this amount may not seem to be a very large percentage of a particular contract, it should be recognized that this cost saving when considered with the extensive operational and maintenance advantages provides an extremely attractive plus factor for any aircraft program - military or commercial.

Conclusion

The CITS under development for the B-1 bomber will provide it with a fast, flexible, accurate means of checkout and fault isolation. The CITS is fast because it is controlled by a digital computer; it is accurate because it uses actual aircraft parameters generated by the onboard systems in their operational environment; it is flexible because its processing of analog data is limited mainly by the software which is much easier and less costly to change than hardware. CITS is, in essence, a growing, living, software-oriented system readily adaptable to aircraft system changes and to incorporation of future technological changes. In addition to its speed, flexibility is one of the CITS most important advantages, since it is this flexibility which provides the test system with much of its cost advantage over BITE and/or other forms of AGE. A CITS type system should provide an aircraft with a mission similar to the B-1 with a cost of ownership advantage of about 12% over a 10-year operational period.

AN INTEGRATED, MODULAR APPROACH TO AUTOMATIC TESTING AND DATA MONITORING

Ernst E. Forster
PRD Electronics, Inc.
Syosset, N. Y.

Introduction

PRD's CAST (Computerized Automatic System Tester) is based upon an integrated system concept. This approach treats the automatic test equipment as a complete working system in an integrated package composed of individual programmable stimulus and measurement equipments which are controlled by a central computer under the direction of test programs and internal software.

Each CAST system is customized to the user's particular requirements by selecting from an inventory of proven hardware and software elements. Requirements for high volume automatic testing, operational maintenance, and data monitoring can all be satisfied by a single CAST system, programmed with a common language and controlled by a common computer.

Both software and hardware are designed in a modular manner so that system expansion can be readily accomplished. In conformance with system modularity, the control and power distribution subsystems are open ended; i.e., they can accommodate the installation of additional programmable instruments ("building blocks") without changing the basic system structure. In addition, the computer's basic capabilities can be readily expanded if operational requirements necessitate.

System Hardware

Figure 1 illustrates the general configuration of the CAST system hardware. The Central Controller consists of a general-purpose digital computer (Typically a DEC PDP-11) and the associated mass memory and peripheral devices. Its primary functions are control and operation of the Stimulus and Measurement Subsystem (SMS), receipt and interpretation of measured data, operator communications and test program compilation.

The Central Controller communicates with the SMS through a parallel interface and a Local Operator Interface which monitors traffic between the Central Controller, Station Operator and the SMS. Functions such as test mode selection, operator instructions and system interrupts are handled through this Local Operator Interface. Communication with the SMS is carried by an open-ended control bus and power distribution system. Each instrument accesses the control bus through dedicated controllers which provide for addressing the individual instruments. This arrangement allows for the control of up to 256 instruments by the single computer Input/Output (I/O) channel.

The instrument complement of the SMS is determined by the user's specific test requirements. An established inventory of instruments with proven CAST compatibility provides a broad spectrum of capabilities:

Measurement

DC Voltage and Current	Synchro/Resolver Parameters
AC Voltage (True RMS)	Waveform Analysis
AC Voltage	Pulse Width
Distortion	Rise and Fall Time
Frequency	Pulse Separation
Period and Events	Pulse Amplitude
Power (RF)	Phase
Modulation	Spectrum Analysis
Logic States Logic Values Propagation Delays	FM Deviation

Stimulus

DC Voltage and Current Sources

AC Single Phase and Multiphase Voltage Sources

Sine, Triangle, and Square Wave Function Generators

Frequency Synthesizers <1 Hz to 18 GHz

Pulse Generator

Precision DC and AC Voltage and Current Sources

Synchro/Resolver Simulators

Multibit Binary Data Patterns

Complex Measurements

Complex parameters which require two or more measurements, units conversion and/or statistical averaging are calculated using the system computer.

System Switching

Signal routing is accomplished automatically through the Programmable Switch. Its unique design provides for programming any contact pin at the UUT interface as either a

digital driver, digital receiver comparator, analog stimulus or analog measurement point with four-wire capability. The Unit Under Unit (UUT) interface is segmented for high frequency (1-500 MHz) and low frequency (<1 MHz) signals and contains both coaxial and standard pin fields.

System Software

The primary role of software in an Automatic Test Equipment (ATE) system is to facilitate efficient communications between the test programmer and the ATE hardware. The test program is the medium through which the test engineer or technician imparts his knowledge of the UUT's test requirements to the ATE system. As such, it must be capable of being written in a language which is easily learned and yet comprehensive to the extent that it fully utilizes the system's hardware capabilities.

The CAST Programming Language

The CAST Programming Language (CPL) is a high-level, English-like language which greatly facilitates the preparation of test programs. It is an ATLAS-based, test-oriented language which permits engineers and technicians with little or no programming experience to prepare reliable test programs. Typically, a test engineer can prepare and debug a CPL test program for a printed circuit board in from 8 to 40 hours, depending, of course, on the board's size and complexity.

CPL statements consist of verbs, nouns, and modifiers which are used to specify the action to be performed. Verbs are action-causing words such as: MEASURE, APPLY, CONNECT, and CALCULATE. Nouns and modifiers are used to direct the verb-caused actions to specific CAST components. The CPL vocabulary also provides for calculating complex parameters from the results of basic measurements and converting between various units of measurement.

CPL Syntax

CPL is composed of statements which enable the user to control each of the various test system elements to generate stimuli and make measurements; perform program branching, looping and iterations; generate messages and data displays on peripheral devices; input data from the keyboard; perform various arithmetic and Boolean operations and evaluate the results of these operations and the data obtained from measurements.

CPL statements follow a general format, consisting of verbs, nouns, modifiers and terminators; this is similar to an English language sentence structure. Specifically, a CPL statement consists of a statement number, a verb, a noun, modifiers, and a statement terminator. Verbs, nouns and modifiers are words which are separated by blanks or commas exactly as in English language syntax. Statement numbers are used to uniquely identify a statement when one statement references another for purposes of branching or looping.

CPL is structured to permit introduction of user-specified variables and data lists for control of the test system. As a result, system stimulus and measurement parameters can be dynamically updated at test program execution time. For example, the user can conveniently step a power supply through a voltage range with just two statements, or he can make a measurement with the digital multimeter, perform

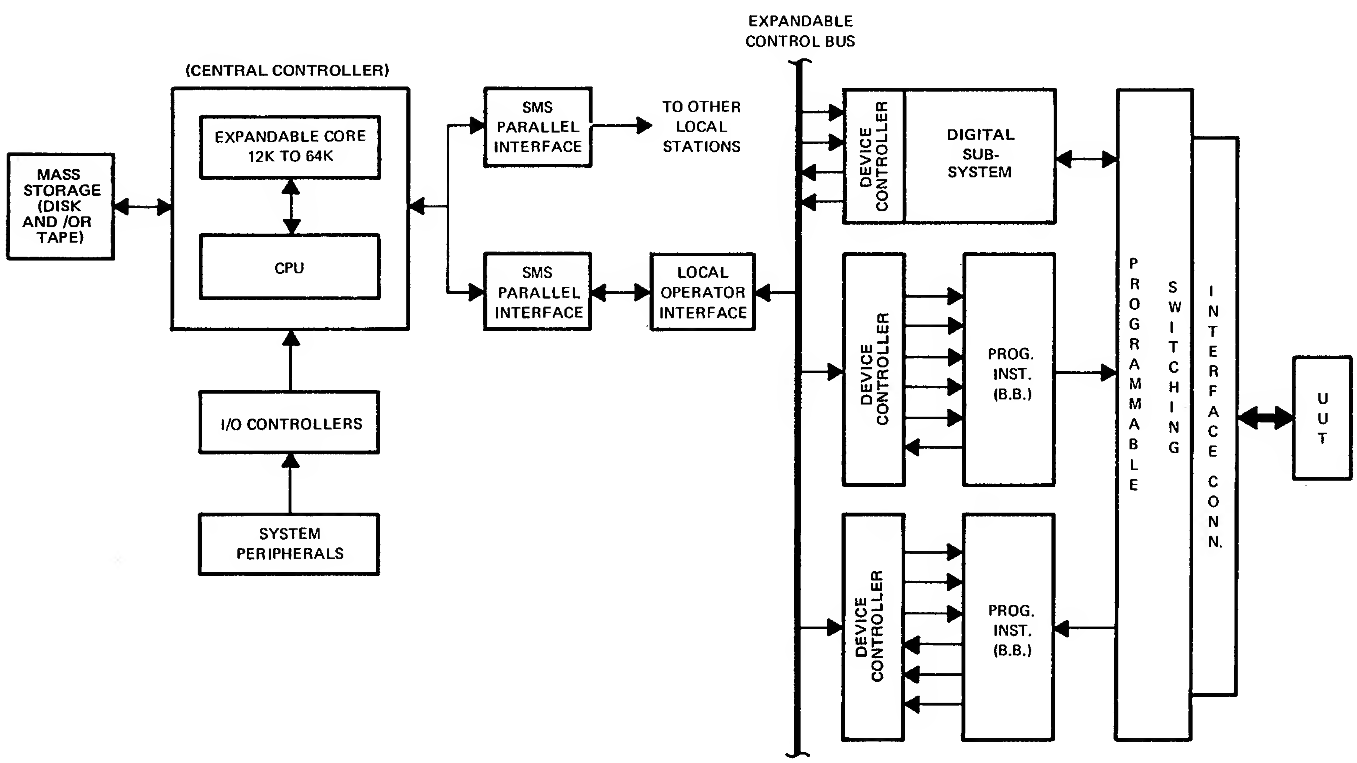


Figure 1

some arithmetic manipulation on the measurement data, and then use the result to set the output voltage of a power supply.

Variables and data lists are referenced by a user-assigned symbolic name which may be up to six characters in length. Variables and data lists may be specified as one of the following data types, depending on the intended use: digital, decimal, integer and character.

Important attributes of CPL are complete programming flexibility, dynamic stimuli control, ease of use, and self-documentation in nature.

CPL Verbs

The following list of verbs with brief descriptions is intended to provide as consistently as possible an insight into the comprehensive nature of CPL:

- BEGIN is used to mask the beginning of a test program and to identify the program for documentation purposes.
- TERMINATE is used to mask the end of all program units. If this verb is executed, test program execution is terminated and control returned to the Multi-Test Executive.
- FINISH is used to halt testing and return all test station components to a quiescent state.
- DECLARE is used to specify data type and allocate storage for user-specified symbolic names which are used to reference variables and data lists.
- DEFINE is used in the test program preambles to specify predefined procedures (subroutines) and messages.
- END is used to mark the end of a preamble procedure. Execution of an END verb provides return out of the procedure.
- WAIT is used to suspend test program execution until one of the following conditions is satisfied:
  - An operator's response is given
  - Manual data is entered by the operator
  - A specified time has elapsed
  - A specified system condition has occurred
- GOTO is used to transfer program control (branch) to any statement in the current program unit. A GOTO statement may be unconditional, computed, conditional or interrupt driven.
- COMPARE is used to perform data comparisons and establish conditions that can be evaluated for the conditional form of the GOTO verb.
- REPEAT is used to perform program iterations conveniently by specifying a sequence of statements to be performed a given number of times. A variable may also be specified which may be incremented each time the iteration is executed.
- CALCULATE is used to perform arithmetic and Boolean operations such as addition, subtraction, multiplication, division, exponentiation, masking and shifting. The CALCULATE statement is similar in structure to that of an algebraic equation. A variable is specified to the left of an equal sign with an expression to the right. Upon execution, the expression is evaluated and the result replaces the current value of the variable. The expression is specified using a combination of variables, list elements, constants, operators and functions.
- FILL is used to place data in previously declared storage locations. FILL may be used in the program preamble section to statically assign constants to variables and list elements. In procedural sections, FILL can be used to input data from a mass storage file or from the input data channel of one of the digital test subsystems.
- PRINT is used to print data and messages on the system printers during test program execution.
- DISPLAY is used to provide data and message displays on the alphanumeric displays during test program execution.
- MOVE is a generalized data transfer verb used to move data from one location (source) to another (destination). The source may be a variable, list element constant or system descriptor. The destination may be a variable, list element or system descriptor.
- CLEAR is used to reset or initialize the various stimuli and measurement devices in a test station.
- CONFIGURE is used to set up various digital test subsystem components as required prior to application of stimuli and collection of response data.
- APPLY is used to generate stimuli using devices such as power supplies or signal generators.
- MEASURE is used to make measurements and prepare the measured data for subsequent evaluation or manipulation.
- ISSUE is used to transmit absolute instructions codes to any of the test system stimuli and measurement devices.

Measurement and stimuli data can be manipulated conveniently using the CPL CALCULATE statement. Arithmetic operations such as addition, subtraction, multiplication, division, square root and exponentiation as well as many others can be used to perform any operations the user requires for a particular test application.

Figure 2 is an example of a program coded in CPL which calls for adjusting a voltage tuned oscillator to a specific frequency.

In this sample program, a voltage control variable, 'VLTFRQ', is first initialized to 0.1 volt. This variable is then used in an APPLY statement to generate a voltage which is used to provide an input signal to a voltage tuned oscillator. The resultant output frequency is measured and if it is less than 100.0 kHz the voltage control variable 'VLTFRQ' is incremented by 0.1. This process continues or loops until the oscillator's output signal reaches 100.0 kHz.

```
     10  BEGIN, VTO ADJUST TEST$
     20  DECLARE, DECIMAL, 'VLTFRQ'$
E        ENTRY POINT $
     100 CALCULATE, 'VLTFRQ' = 0.1$
B        BRANCH OBJECT FROM STEP 140 $
     105 APPLY, DC SIGNAL, VOLTAGE 'VLTFRQ'V, MAXIMUM 1A,
         CNX HI J1A5 LO J1B9$
     110 MEASURE, (FREQUENCY), AC SIGNAL, MAXIMUM 1V,
         CNX A J1C3$
     120 GOTO, STEP 150, IF 'MEASUREMENT' GE 100.0$
     130 CALCULATE 'VLTFRQ' = 'VLTFRQ' + 0.1$
     140 GOTO, STEP 105$
B        BRANCH OBJECT FROM STEP 120$
     150 DISPLAY, C 'REACHED 100.0KHZ WITH INPUT OF',
         'VLTFRQ', C' VOLTS'$
     160 FINISH$
     170 TERMINATE$
```

Figure 2

The CPL Compiler

Programs prepared in CPL source language are compiled on the CAST computer utilizing a 12k core augmented by disk memory. In addition to performing the accepted compiler functions such as sequential processing and storage allocation, the compiler provides for the routing of stimulus and measurement signals through the switching matrix to and from the appropriate pins of the UUT interface, thereby relieving the test programmer of this responsibility. To ensure efficient program operation and error-free generation, the compiler provides the following capabilities:

- Repetition of program sections
- Real-time system parameter control
- Definition and utilization of data arrays
- Report generation
- Reference to previously compiled program sections

The Operating System Program

All phases of system operation, such as compilation, test program validation and debugging, test program execution and self-maintenance, are managed by an operating system program. Included in this program is a monitor which controls the execution of the various operating system elements, such as the loader, interpreter, and executive routines.

The Test Program Editor

A test program editor routine, provided as part of the CAST software, facilitates the printout of test program listings and the debugging and modification of individual test program statements.

System Operation

Program Compilation

When CPL compilation is to be performed, the operator places a previously punched CPL source program tape in the high-speed paper tape reader. The compiler then converts the source program to object code instructions which are interpreted by the computer. When compilation is complete, the resulting object program can be output on paper tape via the teletypewriter punch. The resultant tape may be loaded and executed by CAST at some later time.

The test program can also be compiled and subsequently executed in one operation if the operator so desires. When this mode of operation is requested, the object program is not punched on paper tape but is written in a temporary storage area on the disk. If no errors are encountered during compilation, the object program will be automatically loaded and executed. Any operator responses or intervention required by the test program will then occur only as a result of execution of appropriate CPL statements.

Program Loading and Execution

Program loading is automatic in the CAST System. The object program is called from its file by the simple act of typing "RUN," followed by the test program name. The program is loaded and executed without further operator intervention.

To assure that the requested program is applicable to the UUT connected to the CAST interface, the system scans prescribed pins to read an identifying resistive code built into the UUT. If the code is correct the test is executed; if not, the test is aborted and the operator is so instructed.

The resistive code on the UUT may also be used to call its particular test routine but this mode is not generally recommended since more than a single test routine is usually associated with each UUT.

If the object program is stored on paper tape, the operator places the tape on the reader and types "LOAD" on the keyboard. The program will then run in the mode selected by the mode switches. The following modes of executions are available:

- ONE GROUP - The program halts when a group break is called for in the program and another light will indicate a program halt. The INCREMENT switch is depressed to execute the next group of tests.
- ONE TEST - Execution of the program halts once the designated test is completed and another light will indicate a program halt. The INCREMENT switch is depressed to execute the next test.
- ONE STEP - Execution of the program halts once the designated step is completed and another light will indicate a program halt. The INCREMENT switch is depressed to execute the next step.

- ONE INSTRUCTION - Execution of the program halts as each instruction is completed. The INCREMENT switch is depressed to execute the next step.

- AUTOMATIC - Initiates the automatic running of the complete program.

Consider an amplifier as the Unit Under Test (UUT). Typically, a check of the amplifier's gain and phase shift would be made as a function of frequency.

The ONE GROUP mode of operation would normally contain those test statements necessary to perform a series of amplitude and phase measurements as a function of frequency, whereas the ONE TEST mode would typically involve either amplitude or phase measurements as a function of frequency. Similarly, the ONE STEP mode would involve testing gain or phase at one particular frequency. In addition, the ONE INSTRUCTION mode would typically involve the execution of one computer statement (i.e., CONNECT, MEASURE).

Test results can be printed on the teletypewriter under control of the test program. Any additional printouts or operator directives can also be printed as required.

Digital Test Capabilities

When configured for digital testing, CAST is a modularly structured tester which provides fault isolation as well as functional and dynamic test capabilities. Its modular architecture permits it to be configured to any level of digital testing, from simple combinational logic circuits up to highly sophisticated time-critical modules and systems.

UUT Interface

The interface between CAST and the UUT is available in eight-pin increments from 32 to 128 pins each for both stimulus and response data. Up to 128 of these pins are bidirectional; i.e., they can be programmed as either inputs or outputs. This bidirectional capability permits an entire series of UUT's with mechanically similar connectors to be tested using the same interface device. The standard CAST/UUT interface connector employs TTL-compatible logic and/or programmable logic "1" and logic "0" values.

Independent Multichannel Operation

Independent multichannel operation is employed for transferring stimulus and response data to and from the CAST/UUT interface. Five stimulus data channels provide access to either software or hardware stimulus data generators. Similarly, five response data channels allow either software or hardware processing of response data. This arrangement provides the high-speed data transfer rates generally available only from dedicated testers, while retaining the versatility of a general-purpose test system.

Hardware Data Sources

The availability of dedicated hardware data sources provide:

- Data Transfer Rates in Excess of the Computer I/O Rate

- Conservation of Core Storage

- Simplification of Test Programs

The CAST Programming Language (CPL) provides for initialization of hardware generated data sequences such as:

- Interlaced Rows of Ones and Zeros

- Interlaced Rows of Single and Double Checkerboards

- Walking Bit Patterns

The programmer need define only the initial test pattern, start point and stop point for these patterns.

Digital Test Philosophy

In operation, CAST exercises the UUT at its specified operating speed (up to 10 MHz) and analyzes the resulting response patterns to confirm the truth table requirements. If these requirements are not met, the system program branches to a subroutine which further analyzes the response patterns in order to isolate the fault.

The techniques used in functional testing incorporate a "fault dictionary" which is a computer-stored listing of the failing and passing response patterns for each possible fault in the UUT and the reference designation of the corresponding failing component(s). The size of the fault dictionary is dependent on the complexity of the UUT and the degree to which fault isolation is required. Stimulus patterns, response patterns and fault dictionaries for UUT fault isolation are generated through the use of PRD's Stimulus and Response for Digital Integrated Networks (SARDIN) program.

During the test, response patterns generated by the UUT are compared to reference patterns contained in the test program. Each of the response patterns which fails to pass a comparison is recorded. The failing test numbers are then used to generate a binary fault vector. This UUT fault vector is then compared with the members of the fault dictionary until a pattern match is obtained, which results in a printout of the defective component(s).

Analog Test Capabilities

CAST's analog test capability is limited only by its complement of measurement and stimulus modules. These can be selected from an established inventory of programmable instruments with proven CAST compatibility. Whenever possible measurements of basic quantities (voltage, resistance, current, frequency and time) are used, in conjunction with the computational power of the computer, to determine more complex parameters. This approach minimizes the number of instruments required to meet a user's requirements by taking full advantage of the general-purpose computer. For example, harmonic distortion measurements are made with a digital multimeter and a programmable filter. The signal's fundamental and harmonic frequencies are selectively filtered and measured and the percent distortions for each of all the harmonics is then calculated.

Another example of software enhancement of hardware capability is the "run-time variable." This technique is a software implemented closed loop which permits setting the output of a stimulus instrument to the accuracy of a measurement instrument. For instance, if a 1-milliwatt microwave signal is required to exercise a UUT, the signal generator is

first programmed to 1-milliwatt nominal output. The signal is then measured at the UUT interface by an RF power meter and the generator level is adjusted by the program until the power meter indication is 1 milliwatt. In addition to imparting the higher accuracy of the power meter to the signal generator, the run-time variable also compensates for line losses between the generator and the UUT interface.

Testing a Voltage-Controlled Oscillator (Figures 3 and 4)

A typical analog UUT is a 65- to 117-MHz voltage-controlled oscillator (VCO). Initially, the program checks a resistance signature which is characteristic of the UUT and informs the operator whether the program being run is the correct one for the UUT connected to the interface. Power is then applied to the UUT and the supply current is measured and compared with predetermined limits in the program. Excessive supply currents cause the test to be aborted and the operator to be notified. Increments of tuning voltage are now applied to the oscillator and its output power and frequency are measured at each increment. These measurements are compared with nominal values contained in the program and deviations from nominal are calculated. These deviations are compared with the UUT's specifications to determine test results. The program then requests an output format (graph, chart or go/no-go decision) from the operator and prints the output in this format on the teletype. A typical run-time for this test is 1 minute, excluding printout time which depends on the format selected.

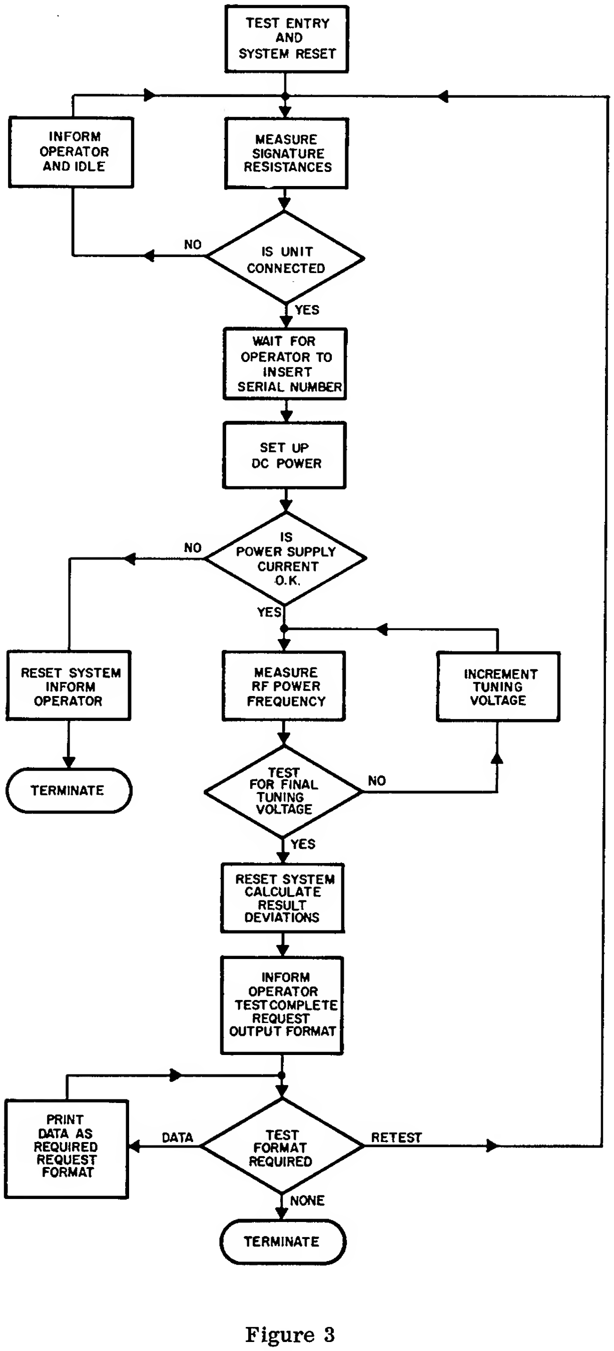


Figure 3

TEST PROCEDURE FOR:

FREQUENCY ELECTRONICS

OSCILLATOR FE31-WCLC

TEST 111

TUNING VOLTAGE	FREQUENCY MHZ	DELTA FREQ.	TEST RESULT	POWER DBM	DELTA POWER	TEST RESULT
-1.0	65.030	1.128	O.K.	29.119	3.814	HIGH
-1.5	68.796	0.446	O.K.	28.957	3.651	HIGH
-2.0	72.292	0.094	O.K.	28.658	3.352	HIGH
-2.5	75.413	-0.633	O.K.	28.381	3.075	HIGH
-3.0	78.608	-1.286	O.K.	28.032	2.726	HIGH
-3.5	82.811	-0.931	O.K.	27.430	2.125	O.K.
-4.0	86.616	-0.974	O.K.	26.862	1.556	O.K.
-4.5	90.157	-1.281	O.K.	26.347	1.041	O.K.
-5.0	94.210	-1.076	O.K.	18.885	-6.420	LOW
-5.5	98.153	-0.981	O.K.	25.288	-0.018	O.K.
-6.0	102.711	-0.271	O.K.	23.823	-1.483	O.K.
-6.5	107.836	1.006	O.K.	21.919	-3.386	LOW
-7.0	111.836	1.158	O.K.	20.527	-4.778	LOW
-7.5	115.653	1.128	O.K.	20.050	-5.255	LOW

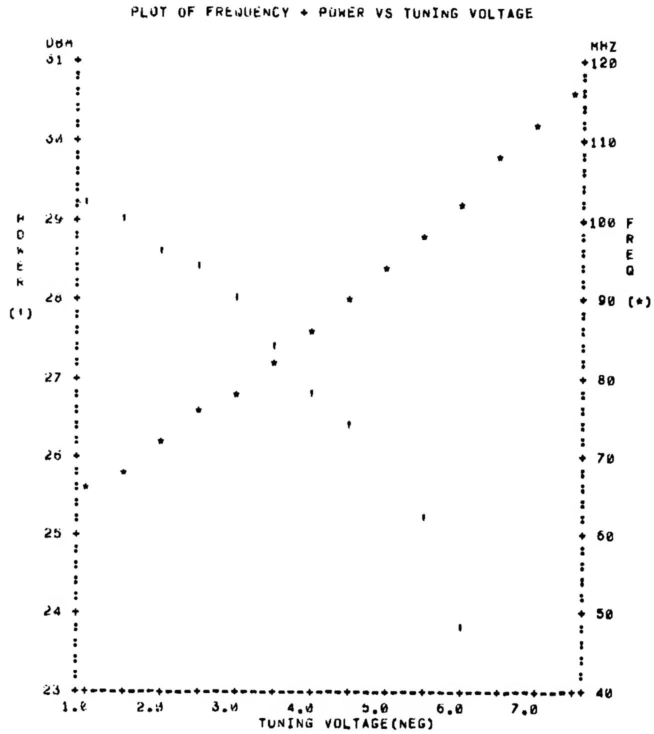


Figure 4

This example illustrates CAST's basic approach to analog testing and the limited demands which it imposes on both the test programmer and the system operator. More complicated UUT's, such as complete transmitters, receivers, modems, multiplexers and their subassemblies are tested by direct extensions of this method.

Fault diagnostics and isolation are performed by including branching statements at evaluation points in the test program. However, as with manual troubleshooting, the level of fault isolation which can be obtained is dependent on the number of test points available on the UUT.

Also, operator interactions, such as alignment and control manipulations, can be called for and the results of these actions monitored on a real-time basis.

Growth Capabilities

Increasing Stimulus/Measurement Capabilities

The basic design concept of CAST was predicated on the requirement that it should be readily modified or expanded to keep pace with the ever changing needs of industry. To add a new test capability requires only that the appropriate stimulus/measurement building block be made available. When a test capability is no longer required, the building blocks associated with that capability can be deleted. All units are rack-mounted and are modular in construction. If all available rack space has been expended, additional racks can be added. A typical system configuration is shown in Figure 5.

Expansion or changes in the stimulus/measurement capabilities has no effect upon the number of I/O channels required or upon the SMS bus. Up to 256 building blocks can be accommodated before any change in the intrasystem communication system would be necessary.

The internal structure of the programmable switch is also modular. If additional switching capability is to be added, it can be accomplished by adding one of the standard switching increments in an expansion section of the switch drawer (four total increments per drawer). Increasing the switching capability does not disturb the existing switches and does not in most cases affect the electrical characteristics of the circuits. Unlike switching matrices, the disjunctive tree switching used in the CAST System does not directly add shunting capacitance as the switch is expanded.

Like CAST hardware, CAST software is also modular. Adding capability does not in any manner disturb or change existing programs. Adding a new building block adds new terms to the programming vocabulary but does not alter the interpretation or limit the use of existing terms. The communication between the computer, SMS controller and the SMS bus is not changed in any way.

Adding building blocks to the CAST System does not alter the self-diagnostic programs. Tests must be added to these routines to service the new units and the total run-time will be increased proportionately, but the existing routine remains intact.

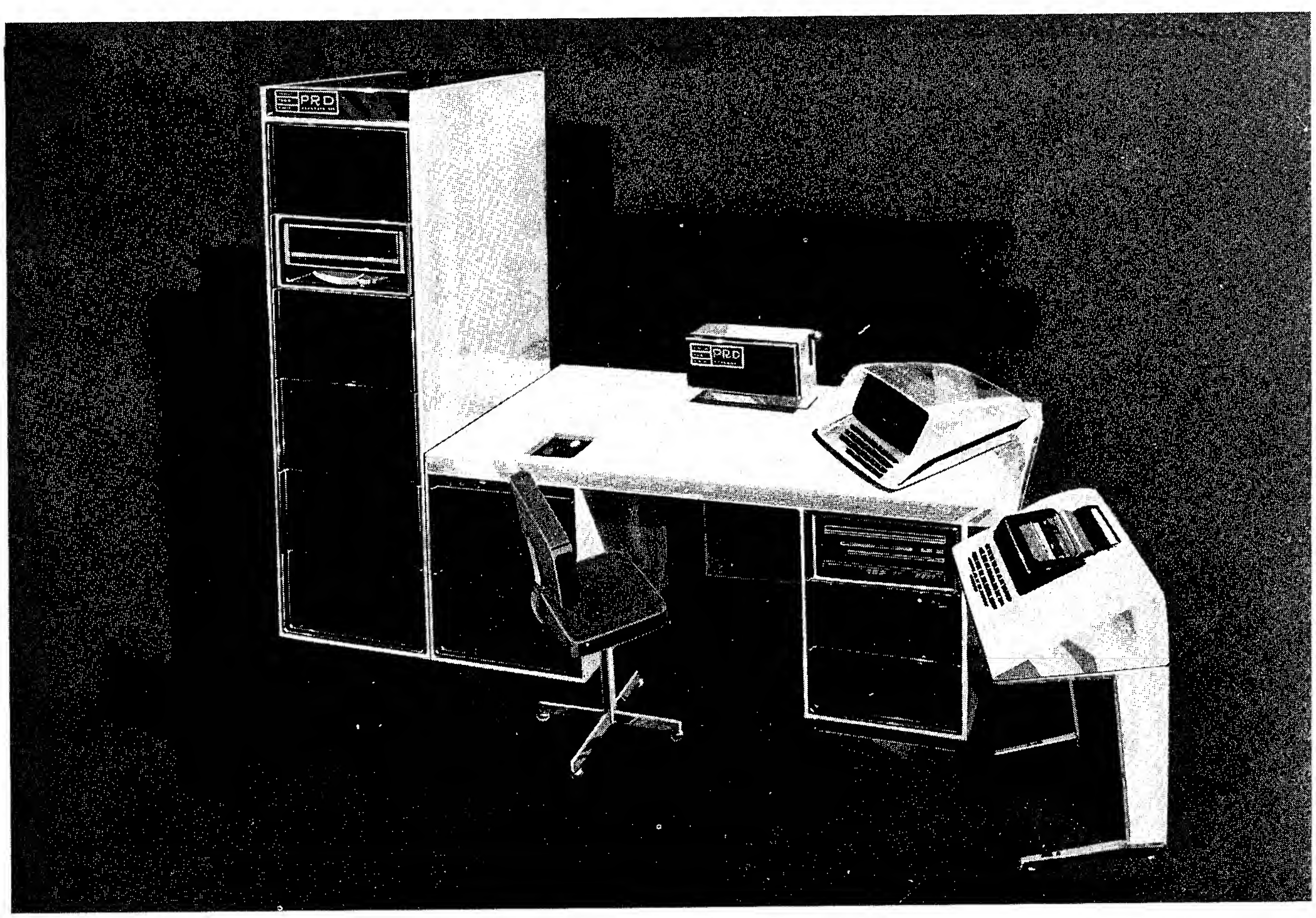


Figure 5.

Increasing Interface Capacity

The interface is modular; containing blocks measuring approximately 3 by 7 inches. Each block contains either 240 low frequency pins or 119 high frequency coaxial connectors, more than adequate for most test applications. However, if a larger number of pins is required to accommodate the testing of large complex units, additional blocks can be added.

Increasing Number of Test Stations

The number of test stations which can be multiplexed is ultimately limited only by the availability of stimulus/measurement building blocks; if a test requires the use of a building block which is tied up in performing a test at another station, the former station must wait. Stations which are performing widely dissimilar test routines seldom place conflicting demands on the available building blocks. Many such stations can be multiplexed without incurring excessive or frequent delays. Conversely, stations running identical routines frequently interfere and suffer delays. An analysis of the anticipated workload is necessary to determine the optimum number of redundant building blocks which should be employed for greatest cost effectiveness.

The CAST software does not limit in any manner the number of test stations which can be operated in a single system. All of the operating routines are capable of servicing almost unlimited numbers of test positions.

The physical length of lines required to service a large number of test positions may make it necessary to place certain building blocks in close proximity to the UUT. As a result, a portion of the switching matrix must be placed at the test station rather than in the CAST mainframe. The modular construction of the switch allows this to be readily accomplished.

In addition to the UUT interface connector provided at each test position, a priority interrupt unit can also be provided. This unit is actuated by the test technician's RUN button or by typing RUN on the keyboard. It encodes an interrupt signal and places it on the computer's interrupt bus where the computer scans the request and establishes its position in the queue.

TOLTS, TOTAL ON-LINE TESTING SYSTEM

by
P. J. Scola
Honeywell Information Systems, Inc.

Phoenix, Arizona

The Series 6000 (see Figure 1) incorporates a new,[1] advanced concept in system availability. The major features of this concept are on-line* test and diagnostic error visibility and preventative and corrective maintenance. Such aids as continuous error detection from the peripheral media to the main memory, programmable hardware margins and subsystem fault registers ensure early fault detection and data integrity in Series 6000 system. On-line test and diagnostic programs operating in multi-programming mode with user programs maximize system availability. All peripherals, communication equipment, main memory and CP's can be tested on-line. Contrast this to a system that does not have on-line test capability and you can readily see why there is a marked increase in system availability to the user (see Figure 2). Automatic retry and recovery on processor and input-output commands is designed to minimize the impact on the operating system of hardware malfunctions. Greatly increased dynamic hardware visibility has been achieved by the inclusion of history registers that record the internal machine states of the last 16 steps performed. These history registers are dynamically snapped and stored for a comprehensive trace of system operation, diagnosis and for retry every time a failure is made or under program or manual call.

The total Series 6000 system is oriented toward optimizing user availability with concurrent maintenance functions. Various portions of the system can be devoted to routine preventive maintenance checks while running user programs. Spot diagnostics can test and diagnose portions of the input-output, communication and central system interlaced with, but not conflicting with, user operation.

All of these advanced features and more, are incorporated in the Series 6000 system to provide the greatest possible system availabil- (see Figure 5). System interruptions to user are minimized by this error recovery and maintainability concept.

As an example, to increase system availability from 95% to 97% you have to increase mean time between system interruption by approximately 50% for a fixed repair time. That's expensive if you use increased reliability to get there. Typical plot of availability curve vs. MTBF shows the reason why this is so (see Figure 3).

Any examination of a computer system and its failures will indicate peripheral units produce the most failures and they can be masked off easily by on-line testing. Items like the central units, memories and communications take special hardware aids to put them under "on-line" testing (see Figure 5).

The total maintenance and recovery concept is integrally incorporated with the operating system. The Total On-Line Testing System (TOLTS) (see Figure 4) is composed of four major subsystems. These are: Peripheral On-Line Testing System, Communications On-Line Testing System, Central System On-Line Testing System and Remote On-Line Testing System. This Total On-Line Testing concept is a Honeywell first[1] in the computer business; it permits up to 24 concurrent diagnostic programs to be operating with user programs in the Series 6000 system.

Some of the major benefits of the Total On-Line Testing System (TOLTS) are:

- The test system provides a complete library of comprehensive on-line "test pages"** designed especially for each system module.
- It provides manual "test pages." These permit the maintenance engineer to design and execute his own test programs using the conversational Test and Diagnostic Language - concurrent with the user's operations.
- The test system is called in automatically or manually from system mass storage.
- 24 test and diagnostic "test pages" can be run concurrently with user operations.
- It limits the amount of main memory used by dynamically allocating and releasing memory. Only the required amount is used.
- Master, auxiliary or remote consoles can be used to call TOLTS into execution.
- All operational and error messages for the "test page" are directed back to the console that initiated the original request and to any other console for monitoring.
- Copies of the error messages can be directed to the system logging file. Or, by completely bypassing the console, messages may be used and accumulated on the file for later analysis on demand.
- OS and TOLTS monitor all error status signals and notify the user of malfunction on a dynamic basis. Error thresholds are set; when they are exceeded, TOLTS can automatically request test and diagnostic assistance or optionally print a message to the operator on the system console. This permits the rapid call-in of the appropriate on-line test and diagnostic program for further fault isolation.

*On-line: The ability to perform the total maintenance function in multi-programming mode while customer is still using system.

**Test page: a collection of tests for a single unit or device.

An additional advantage is achieved by the Test and Diagnostic System in that tests are executed in the same environment as the user programs. Since TOLTS is on-line and an integral part of the total operating system, the user is able to establish a higher equipment confidence level.

The system reconfiguration capability permits any processor to become the control processor. This will permit an easy way to graceful degradation in a redundant system configuration without loss of user operation.

The on-line T&D system is divided into the following five privileged slave programs (see Figure 4):

1. TOLTS Executive.
2. MOLTS (Mainframe On-Line Testing System).
3. COLTS (Communication On-Line Testing System).
4. POLTS (Peripheral On-Line Testing System).
5. ROLTS (Remote On-Line Testing System).

Any or all of these programs (except the TOLTS Executive) may be individually swapped out of core by OS. A single copy of the TOLTS Executive and each subsystem executive is capable of (within reasonable limits) simultaneous execution of any combination of T&D programs controlled from any combination of local consoles or remote terminals without restriction as to which T&D programs are controlled by which terminals.

The TOLTS Executive handles all communication between T&D subsystems and the operator or OS modules (see Figure 6). The TOLTS Executive may be called and/or have entries placed in its execution queue automatically by OS modules, or by the operator from a local or remote terminal. T&D subsystem executives are spawned and controlled by the TOLTS Executive.

All messages from T&D subsystems are passed to the TOLTS Executive which then passes the messages on to the destination (e.g. local console, remote terminal, dedicated printer, etc.)

The TOLTS Executive spawns T&D subsystems as required with "PRIVITY" granted by the OS.

The TOLTS Executive is capable of buffering error messages transmitted to it by T&D subsystems. This is accomplished by means of a rotating list of message specifiers in the TOLTS Executive which are used for dynamic allocation and deallocation of message area within a message data block. The message data block is 320 words large. This allows buffering of at least two long messages and up to 20 short ones. This mechanism results in operation which does not hold up test program execution because of error message outputting except in situations where a heavy column of error message output is occurring. For example, it should be possible to buffer 20 or more images in the TOLTS Executive without holding up a test except for the time taken to move error message images to the TOLTS Executive buffer area.

Error messages must be transmitted by subsystem executives to the TOLTS Executive. It is necessary for the subsystem executive to make itself unswappable and unmovable, access a gated table to allocate buffer space in the TOLTS Executive, move data into the buffer area, update a message specifier to cause the TOLTS Executive to output the message and then enable the TOLTS Executive.

To request that "TOLTS" output or output/input a T&D message, a call sequence must be used in master mode,with system index registers set to their correct OS master mode conventional values.

This call to TOLTS acts as a request to move the associated data from the requesting program to available buffers in TOLTS to be later sent to the appropriate terminals. Inasmuch as there may not be any available buffer space in TOLTS to place the data, TOLTS may not be able to handle the request. For this reason, the first instruction after the calling sequence will be used as a "denial" return to where return will be made if the data cannot be moved from the user program into TOLTS. If the data can be moved into TOLTS, the return will be made to the "acceptance" return, which immediately follows the "denial" return in the call sequence.

Whenever a write (no read) action has been completed by TOLTS, the TOLTS Executive will place an entry into the individual test subsystem's input queue.

Whenever a requested read action has been completed by TOLTS after the data has actually been read into a buffer in TOLTS, the TOLTS Executive will place an entry into the individual test subsystem's input queue.

Following the placing of this entry into the subsystem input queue, TOLTS will cause a dispatch to the subsystem concerned to wake it up.

The test subsystem must obtain the data read using a call sequence.

If, after a TOLTS I/O request, a return is made via a "DENIAL" address return, no data will have been moved into TOLTS. The requesting subsystem must not issue any further requests until TOLTS can free buffers for the requested I/O. Whenever the denial return must be taken after the request has been made an entry will have been placed into a buffer in TOLTS so as to reserve all subsequent buffer space for the denied request. TOLTS will monitor all buffer spaced released and whenever the denied request can be serviced, TOLTS will place an entry into the particular test subsystem's input queue.

After placing this entry into the subsystem's input queue, TOLTS will enable the test subsystem through the OS dispatcher. It is expected that the test subsystem will now repeat the denied request.

TOLTS will be capable of handling one denied request for each test subsystem (three at present), and will queue them in priority order in the order in which the denial is given. If, after one test subsystem has been given a denial return, a second subsystem issues a request, that second subsystem will unconditionally be given a denial so that the denied request for the first subsystem can be given priority.

T&D SUBSYSTEM ORGANIZATION

The subsystems described in the following sections will have the following features:

1. Each subsystem is executed as a privileged slave program.
2. Each subsystem resides in a contiguous segment of core containing Subsystem Executive (MOLTS, COLTS,

or POLTS), followed by the test programs.

3. Each subsystem executive has a task dispatcher.
4. The subsystem executives will issue calls to acquire more memory in order to load test programs which run under their control.
5. The subsystem executives will be responsible for memory compaction and issue calls to release memory when a test terminates.
6. Infrequently used sections of code are segmented.
7. Operator selected test sequencing capability is provided by each subsystem.
8. A common subset of options is included in each subsystem.
9. All messages are outputted or inputted via the TOLTS Executive.
10. The POLTS Executive will interpretively execute individual peripheral tests. The MOLTS and COLTS Executives will execute individual tests under assembler language control.

To facilitate intermodule communication, a common set of definitions is used by all executives (including TOLTS) and test pages to define those location values used for this communication. This common set of definitions consists of a program skeleton (in source format on a tape) with the appropriate symbols defined and commonly used macro-skeletons. The second file on this tape will consist of a duplicate of the above definitions and the common coding for the subsystem executives with appropriate conditional assembly of those functions unique to a particular subsystem.

The following is a general memory layout for the subsystem executives:

a. Entry points.
b. Common coding for executive functions.
c. Specific coding for executives.
d. Common data conversion routines.
e. Common message setup and output routines.
f. Specific executive message setup routines.
g. Constants and special tables.
h. Initialization coding, to be overlayed by the first test page call.

The common coding of the subsystem executive will process special requests or calls from the test programs for routines that are part of the OS:

1. Stick in program until all outstanding I/O is completed.
2. Release resources from this program.
3. Snapshot dump memory.
4. Terminate and release all resources.
5. Abort and release all resources.
6. Get date and time.
7. Set Loop time limits.
8. Calls in program after time delay.
9. Set memory bounds to smaller limits within allocated bounds.
10. Bypass program for execution.
11. Etcetera.

The following option characters will be processed by the common coding in the subsystem executives:

A - Accumulate the error messages on the statistical collection file (test page start and term messages unconditionally go there).
B - Bypass error message output.
C - Give details of all errors including dump of all words in error, etc.
E - Output transient error messages.
H - Halt for input of options following error messages, test end message, pass end messages and cycle end messages.
I - Inform operator of test end.
L - Loop on current test (cannot loop on test 0).
N - Negate the following option character.
O - Go to "ENTER OPTIONS" following complete processing of the current option string.
P - Issue an end pass message any time a back jump is detected while sequencing through tests.
R - Issue an end cycle message any time the test page would normally end and recycle back to start the page again.
S - Unconditionally skip to the next test.
Txx - Unconditionally jump to start the test specified.
Z - Trace; this option is to be used for debug and each test page can use it as a flag to output snap dumps, etc. Any other character will be considered illegal.

The following control mnemonics (.OPTIONS) will be processed by the common coding in the subsystem executives only if found at the beginning of the option string:

.GO - Return to the test page where interrupted unless "S" or "Txx" has been specified. Next test selection will be done for the latter two cases.
.OPT - An "ENTER OPTIONS" message will result immediately.
..PR2 - A request for a dedicated printer will be made to TOLTS unless it is already available. If a regular assigned printer has been allocated, the request will be denied; when the printer is available, the printer available flag will be set.
.PRT - A request for an allocated printer will be made to TOLTS unless it is already available. If a dedicated printer has already been allocated, the request will be denied; when the printer is available, the printer available falg will be set.
.TYP - The printer will be released for the test page if it had been requested and all future messages will be put on the controlling console/terminal.

- .TAL - A tally of all errors will be output.
- .TEST E - The test page will be force terminated.
- .TEST W - The subsystem executive and all test pages executing under that executive will be force terminated (wrapped up).
- .WAIT - The test page will be put in a waiting condition and a "WAITING" message will be output every minute.
- .SEQT - The test table will be sequenced to the users ordering within the following restrictions:

  a. The test numbers must be one to 80 digits separated by commas.
  b. Any test number cannot be zero and must lie in the current segment.
  c. The test table size cannot be increased. No more than the original number of tests in the segment can be specified.
  d. A minus sign (-) preceeding a test number is allowed and indicates a jump (either forward or backward) to the first occurrence of that test number in the new sequence.
- .SEQR - The test page will have its test table resequenced to its original value.

If an .OPTION is encountered which does not match this list, then a check will be made to determine if the test page can process other .OPTIONS. If not then the input is illegal.

MOLTS (Mainframe On-Line Testing System)

MOLTS includes the MOLTS Executive and those programs which executive under its control.

The MOLTS Executive consists of a loader, a task dispatcher, a fault handler to process faults which occur during execution of test programs, and an option processor.

Main memory storage modules, system controller modules, control processor modules, input/output channels, and input/output multiplexer modules can be allocated to the Total On-Line Testing System concurrently with user operation. Now for the first time, a large-scale multi-programming, multi-processing system can be maintained with minimal off-line maintenance.

COLTS (Communications On-Line Testing System)

This new extension of the TOLTS permits test and diagnostic programs to be run on all DATANET* 305's, DATANET* 30's, DATANET* 355's, High-Speed Line Adapters, Low-Speed Line Adapters, teletypewriters, DATANET* 355 card readers, DATANET*355 consoles and DATANET*355 GERTS input/output systems. Again, these tests can be under either local console or remote teletypewriter control. This Communications On-Line Testing System opens a new vista to system maintainability and availability.

*Trademark

POLTS (Peripheral On-Line Testing System)

OS includes a comprehensive Peripheral On-Line Testing System, comprising an executive and a set of test and diagnostic routines. Special interfaces enable the diagnostic testing of peripheral devices concurrent with the production workload. Furthermore, OS accumulates recovered error statistics for continual measurement of peripheral device performance. Through subsequent analysis, problems can be detected and corrected before they become critical.

ROLTS (Remote On-Line Testing System)

For those systems that have remote terminal capability, TOLTS provides the ability to use a remote teletypewriter terminal as if it were a local system console. For those problems that require a maintenance specialist, it will no longer be required to wait for the specialist to travel to the malfunctioning site. Instead, by using a standart teletypewriter and the telephone network, the specialist can dial into the computer system and be automatically connected to TOLTS. The specialist will then have the full range of operating features of TOLTS programs plus his own designed programs available to him. All error messages for the module test will be directed to the local console for the site maintenance engineer, and to the remote teletypewriter for the maintenance specialist, with the additional ability to transmit copies of the TOLTS messages to still other teletypewriters for monitoring purposes. By getting firsthand knowledge about the malfunction via remote TOLTS, the specialist will be able to instruct the site maintenance engineer as to the corrective action to be taken. By resolving the problem in this manner, system down-time will be considerably reduced since the malfunctioning module is out of service for a shorter period of time. TOLTS provides the maintenance engineer with the capability of accumulating all TOLTS error messages on a dedicated system accounting file. The accumulation of error messages and related diagnostic data can be made available to the maintenance specialist via the remote teletypewriter console. This advanced system concept is the result of the continuing evolutionary maintenance techniques developed on the Honeywell Series 6000 systems.

OPERATOR COMMAND STRUCTURE

The following items were considered in the command structure design:

1. There should be as little impact as possible on present OPTS-600[1] options.
2. The command structure should allow addition of new commands and options without requiring any change to the initial commands.
3. The command structure should be as clean, simple and easy to use as possible.
4. The command structure should be

standardized between subsystems to the extent feasible.

5. The primary function of the TOLTS Executive should be to pass the command on to the appropriate T&D Subsystem rather than to process the command.

The command structure is defined as follows:

TESTb̸cccccc

or

TESTb̸xyyyyyyyyyyd

where,

TESTb̸ - is the character string used to specify to OS that the command is for TOLTS.

cccccc - is one of the following TOLTS system commands:

BYE	Orderly disconnect from TOLTS (remote terminals only).
COPYxx	Copy test page output from terminal xx.
NCPYxx	Cancel copying test page output from terminal xx.
LSTAL	List all test pages (active or queued) on the system (passed to subsystem executives).
W	Wrapup all TOLTS operations on the system.

The TOLTS Executive processes the command directly if it is one of the system commands listed above. Otherwise, the TOLTS Executive will pass the command on to MOLTS, COLTS, or POLTS depending on the first character of the command. If the command is not one of the system commands listed above, it must be formatted as follows: (see Figure 3)

TESTb̸xyyyyyyyyyyd

Where,

x - specifies the T&D subsystem as follows:

- M - MOLTS, Mainframe On-Line Test Subsystem.
- C - COLTS, Communications On-Line Test Subsystem.
- P - POLTS, Peripheral On-Line Test Subsystem.
- B - POLTS, BOS (POLTS driving BOS 355).

yyyyyyyyyy - This variable length string of (up to 10) characters is the T&D subsystem command string which is passed by the TOLTS Executive on to the specified T&D subsystem executive. The format of this string is described below for each subsystem. However, for current subsystems the yyyyyyyyyy string is subdivided into a 1 character ACTION field (null on initial test page call, 0 - enter options, or E - end test page) followed by a variable length IDENTIFICATION field, followed by OPTION characters.

d - This character is appended to the string as typed by the program which reads the command from the terminal. It contains the coded ID of the terminal used by the operator. The TOLTS Executive keeps a table of the actual ID based on this code.

ERROR MESSAGES

Each TOLTS subsystem (MOLTS, COLTS and POLTS) has its own error message formats. However, all error message formats are standardized to the extent feasible.

POLTS error messages are similar to the current OPTS-6000 error messages. MOLTS error message formats are defined, and COLTS error messages are similar to POLTS error messages except where there is a good reason to deviate.

The test program initiating the error message may (and normally will) append characters onto the standard "left part." For example, an LSLA test program may append the following data as the "right part" of the first line:

```
                                          SUBTEST NUMBER
XX Command Description INT:
      ee/aa  STAT:ee/aa,DATA ERRORS:xxxx
Expected/actual number of interrupts
Expected/actual status
Number of data errors in this test
```

In some cases where all of the information desired will not fit in the first line, a second line may be required. Thus, Part 1 consists of 1 or 2 lines of information. Part 2 is expected to be a few lines (normally 1 or 2) of prose diagnostic or information statements (e.g., "PROBABLE CAUSE IS SLA BOARD").

Part 3 is detailed additional information which is outputted only if the D (Detail) option is set. Also, if operating in the halt after error mode, part 3 of the last error message will be outputted upon reissue of error message by means of the .HELP command.

REFERENCES

1. Nelson, G.W., OPTS-600 On-Line Peripheral Test System, Fall Joint Computer Conference, 1968.
2. Scola, P.J., Increasing System Availability by Using On-Line Test and Diagnosing, IEEE Int. Symposium Fault-Tolerant Computer, Pasadena, Calif., 1971.
3. Scola, P.J., Logic Design Guidelines for Maintainability, IEEE Int. Convention, New York, N.Y., 1972.
4. Scola, P.J., On-Line Computer Testing System, IEEE Workshop, Lehigh University, Bethlehem, Pa., 1971.
5. Scola, P.J., System Self Repair, Proc. 1971 Computer Design Conference, Anaheim, Calif., 1971.
6. Scola, P.J., Annotated T&D Bibliography, Honeywell Computer Journal, Vol. 5, No. 4, 1972

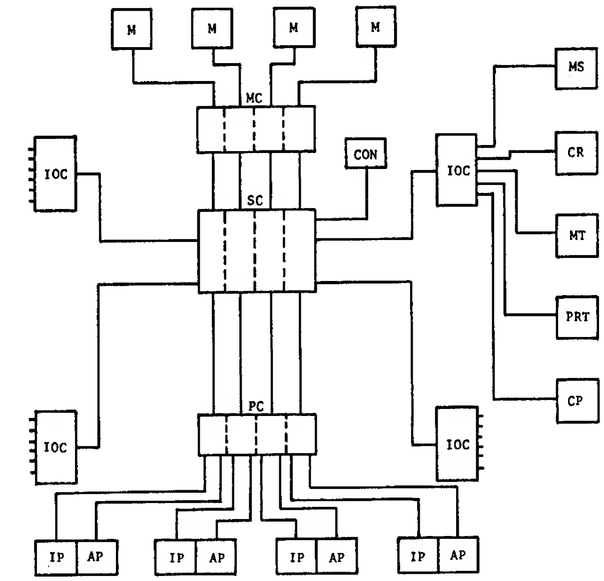


FIGURE 1

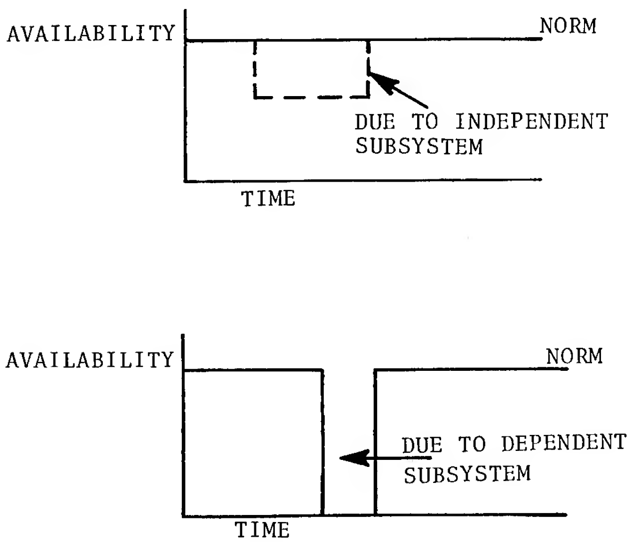


FIGURE 2

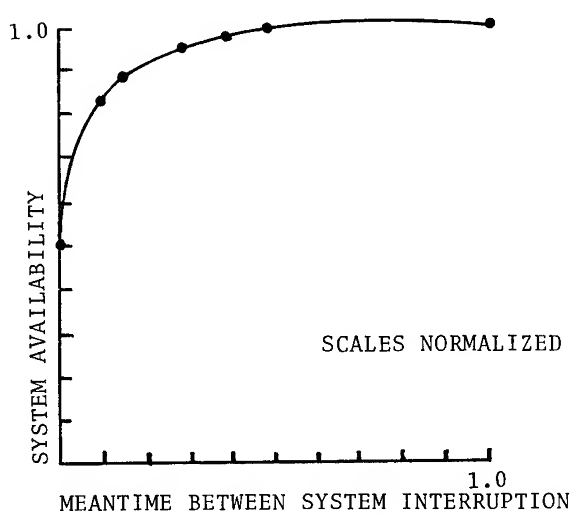


FIGURE 3

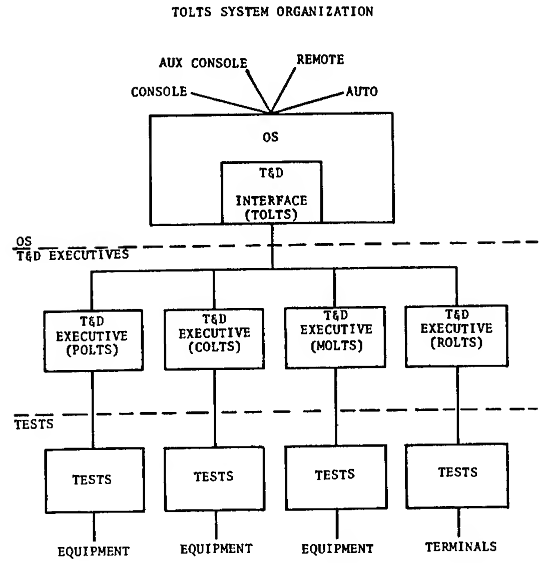


FIGURE 4

1. History registers - snapshot registers
2. Instruction retry
3. Mode register
4. Status register (programs and boxes)
5. Fault register (error)
6. Integrity checks
   a. Data
   b. Address
   c. Op code
   d. Interface
   e. Adders
7. Correction codes
8. Multichannel
9. Dynamic maintenance panel
10. Wraparound
11. Instruction step and save
12. Controlled traps
13. ROM T&D
14. External diagnostic processor for decision making and error logging
15. Multi bus
16. Logic partitioning
17. Diagnostic commands
18. Programmable margins
19. Box utilization or reset
20. Command simulation
21. System and subsystem organization
22. Masking

FIGURE 5

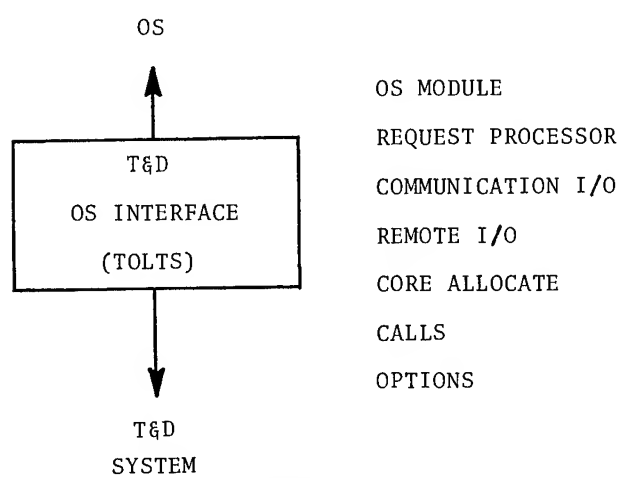


FIGURE 6

"TEST M X YY ZZ ØØØØ"

VERB

TEST SYSTEM

I/O CONTROLLER

CHANNEL/ADAPTER

DEVICE/TERMINAL

OPTIONS

EXAMPLE

"TEST PO1902IR"

WHEN

TEST:	CALL T&D SYSTEM
P:	PERIPHERAL TESTING
0:	IO CONTROLLER #0
19:	DISC CONTROLLER ON IO CONTROLLER CHANNEL #19
02:	DISC DEVICE #02
I:	INDICATE ON CONSOLE EACH SUB TEST START
R:	RECYCLE TEST WHEN DONE

FIGURE 7

INDEX SERIAL NUMBER – 1085

CONCEPT AND SYSTEM OF THE VERSATILE AVIONIC SHOP TEST (VAST) SYSTEM

O.L. Eichna, Jr.
PRD Electronics, Inc.

In the latter part of the 1950's the Navy became increasingly aware of, and concerned about, the problems associated with the maintenance of airborne electronic (avionic) systems in the carrier environment. The problems generally cited were the cost of the maintenance equipment, lack of shop work space, lack of sufficient capable technical personnel, and support of the maintenance equipment itself. Furthermore, the trends in all of these areas indicated that the problems would grow more severe if remedies were not developed and implemented.

PRD Electronics, Inc., undertook a study program with the Navy in 1960 to develop and recommend a maintenance philosophy which would deal with these and other maintenance problems. As part of this effort, the Navy's procurement practices for maintenance and avionic equipment were studied, carrier maintenance shop activities were evaluated during actual operations, and the Navy's maintainability specifications were evaluated.

The study indicated that the primary cause of the maintenance problems was the fact that each avionic system and aircraft had its own special maintenance support requirement. Navy maintenance philosophy at that time was oriented about special support equipment for each aircraft. This alone was enough to cause the problems encountered.

Under this conceptual approach each aircraft had a unique set of maintenance equipment, unique personnel training and skill requirements, unique logistic support for the maintenance equipment, and a unique maintenance management team as depicted in Figure 1. Support equipment for one aircraft was rarely usable for another. Because of this, the amount of support equipment which had to be in the shop increased each time that a new aircraft was added to the carrier complement. This fact was the prime cause of the work space problem.

This wide variety of maintenance equipment and techniques also aggravated the problems associated with personnel training and turnover. The technician was faced with knowing and understanding most of the operation and repair of all the maintenance equipment; the Navy was faced with training him, and then his replacement after the technician's short service period was over. Training costs were therefore high and personnel with the high learning capacity required to assimilate the extensive knowledge were in short supply.

Since the support equipment itself was considered to be special to an aircraft, its development tended not to utilize designs developed for other support equipment and, in many cases, the wheel was reinvented several times. Costs for the equipment were therefore relatively high for the real value obtained. In addition, this approach led to unique spare parts for each equipment, not only increasing spares costs but also burdening the supply lines and storerooms.

Some attempts were made to alleviate the problems in the shop by increasing the testing and repair on the flight deck. This approach was doomed to failure because it simply transferred the problems to the flight deck in the form of a number of unique "suitcase" testers. These testers soon cluttered the flight deck to the point where the area was unable to perform its primary function, preparing a plane for flight.

As a result of a careful study of the problems described, PRD made two basic recommendations: (1) implement Built-in Test Equipment (BITE) in the aircraft to isolate failures to a Weapon Replaceable Assembly (WRA) (2) develop standardized test systems for further fault isolation and repair in the maintenance shop. In addition, three specifications were generated which would enforce compliance with this approach. One specification defined the capabilities of the tester: avionics would have to be designed to be maintained utilizing this tester. Another specification served to ensure that adequate and meaningful test points were incorporated into the avionics. The third specification defined the techniques for using the tester.

The tester developed is the AN/USM-247(V) Versatile Avionic Shop Test (VAST) System shown in Figure 2. In order to define the basic electrical test capabilities in VAST, the test requirements of more than one hundred avionic WRA's and their subassemblies were examined and tabulated. The data was then correlated in a logical fashion so that conclusions could be drawn with respect to test capability. For example, in the area of DC power the voltage was plotted as a function of accuracy, current, and resolution. For signal source requirements, similar plots were developed for frequency, power, and modulation. This information also served to highlight test requirement trends and, together with an evaluation of planned Navy development efforts, led to revisions in the basic data to account for future requirements in order to avoid rapid obsolescence of the test equipment.

During the study of the Navy's maintenance problems, it was observed that a particular piece of special support equipment often contained several functional elements (e.g., DC voltage generation, frequency measurement, etc.). Even though each of these special support equipment units could perform more functions than were utilized, they were limited by the specialized and fixed interconnections within the equipment. This information and evaluation led to a VAST System requirement for modularity and served to define the level of the modularity. The VAST modules, or building blocks, each contain basic electrical functions that, based upon the test requirement study, were not required simultaneously. In addition, each building block was to perform its basic function independently of other building blocks. Not surprisingly, the basic functions of the building blocks are similar in many respects to commercially available laboratory test equipment or

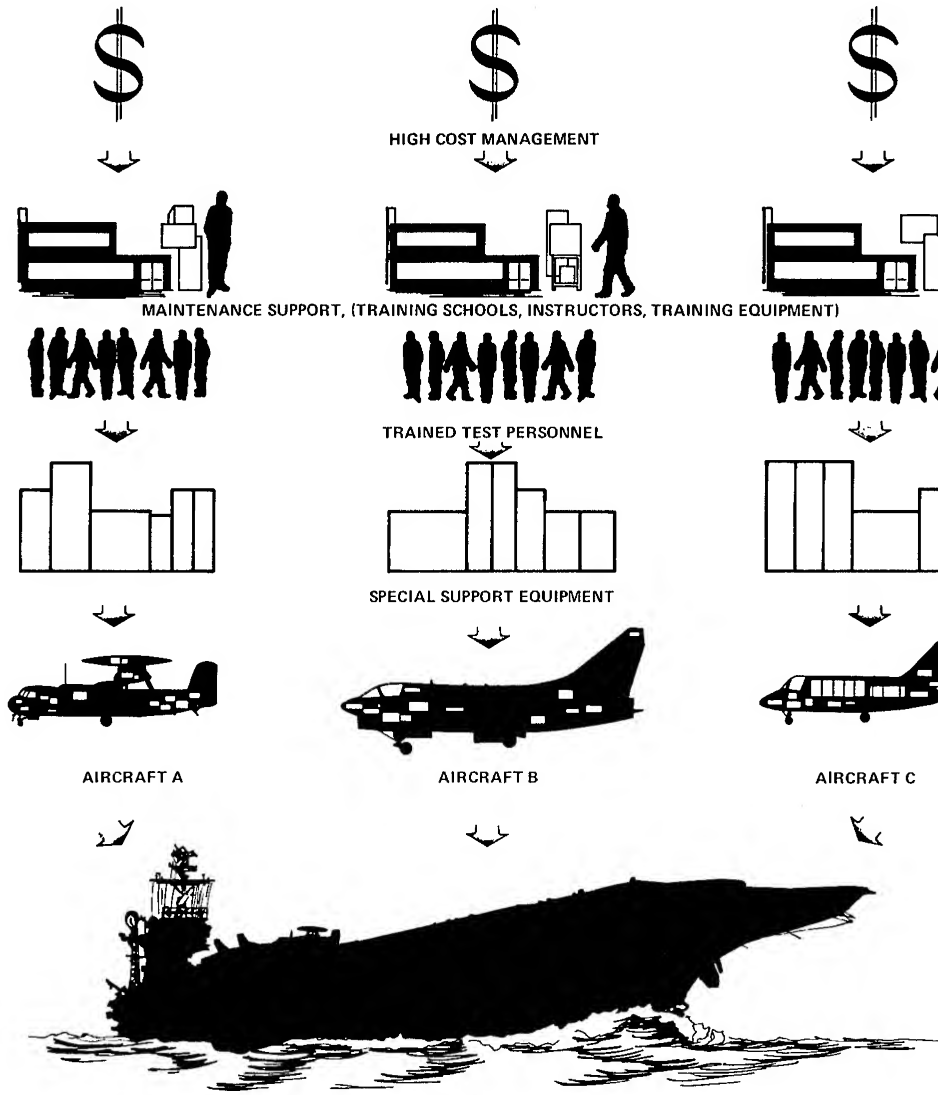


Figure 1

programmable instruments. The building block complement includes power supplies, signal generators, voltmeters, counters, etc. See Figure 3.

Having defined the basic building block functions and the test capability of VAST, a set of building block specifications was developed which apportioned the system test capability to the appropriate building blocks. In addition to the basic functional building blocks, a system switch was also defined based initially upon the test requirements. This switch serves to carry signals between the Unit Under Test (UUT) and the proper VAST building block. It also interconnects VAST building blocks to provide more complex electrical functions and to allow building blocks to test one another in the system.

Since one of the initial concerns was the personnel skill requirements, the system is under computer control and needs little operator participation in performing tests to diagnose failures. In addition, operator interpretive error is practically eliminated. Automatic controllers other than a computer were evluated but these all tended to compromise the intended flexibility of the test system which was inherent in the building block concept. The computer is accompanied by peripheral equipment which serves to load test programs, store executive routines, provide printouts and assist in maintenance.

A Magnetic Tape Transport Unit (MTTU) serves to "read" the test program into the computer for execution. Magnetic tapes were chosen instead of disks due to the service environment experienced on carriers. Prior to compilation for use with the system, the test program is written in a programming language entitled the VAST Interface Test Application Language (VITAL). The language itself has been evolved over the years. Initial versions were quite low-level languages and approached computer code. Recognizing the impact of this language complexity upon the cost of generating test programs, the level of the language has been continuously elevated. At the present time the test programmer can write in normal, test-oriented terminology such as shown in Figure 4. Furthermore, when specifying the points at which a signal is to be applied or measured, the programmer defines them in terms of the UUT nomenclature; the compiler automatically determines and implements the signal path and accounts for path losses. However, if a language is only elevated, the full flexibility of the hardware cannot normally be

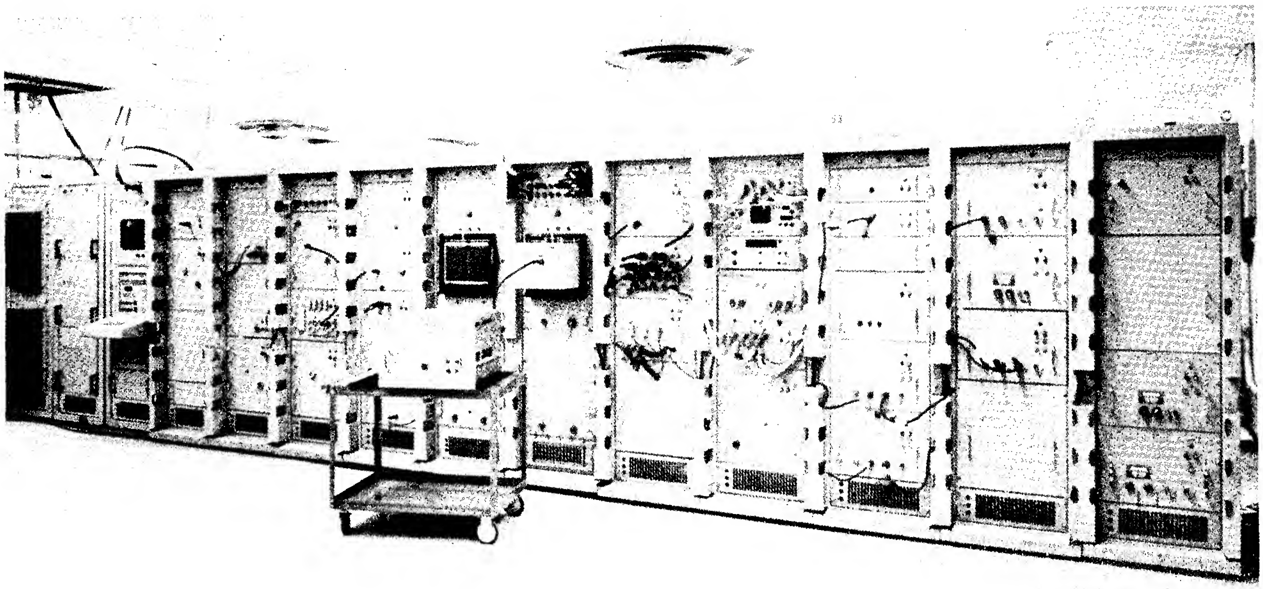


Figure 2

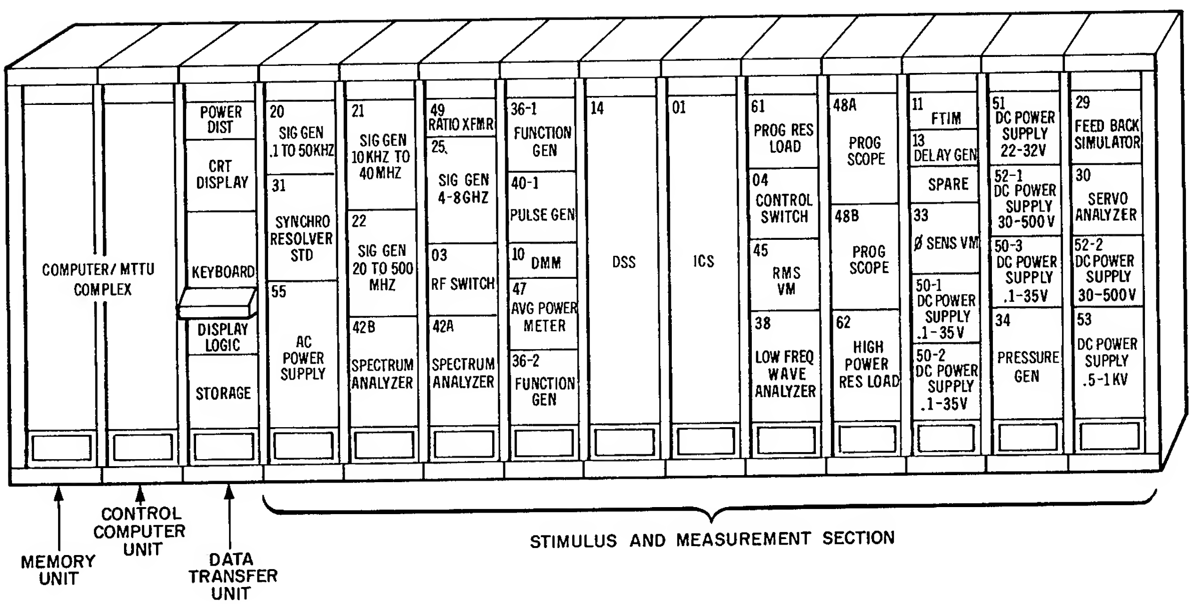


Figure 3

preserved because the higher order language only implements a limited number of combinations. Recognizing and countering this limitation, VITAL deliveretely contains within its vocabulary all of the lower order language elements to take advantage of the full hardware capability. The lowest level allows the test programmer to define and transmit a single digital command to a building block.

This arrangement minimizes the need for highly skilled test programmers. Most of the programs can be written by programmers in the straightforward, higher level language.

The modularity of the hardware is carried through in the computer software. Information unique to a building block, (e.g., range of operation, digital command formats, etc.) necessary for compiler operation is contained in tables rather than being incorporated in the main compiler functional flow. This allows for modification of existing building blocks or the addition of new ones (with associated programming language changes) without affecting the fundamental compiler design.

To obtain a test program tape, the program is written on standard punched cards. The card deck accesses the

```
     10 BEGIN, VTO ADJUST TEST$
     20 DECLARE, DECIMAL, 'VLTFRQ'$
E       ENTRY POINT $
    100 CALCULATE, 'VLTFRQ' = 0.1$
B       BRANCH OBJECT FROM STEP 140 $
    105 APPLY, DC SIGNAL, VOLTAGE 'VLTFRQ'V, MAXIMUM 1A,
        CNX HI J1A5 LO J1B9$
    110 MEASURE, (FREQUENCY), AC SIGNAL, MAXIMUM 1V,
        CNX A J1C3$
    120 GOTO, STEP 150, IF 'MEASUREMENT' GE 100.0$
    130 CALCULATE 'VLTFRQ' = 'VLTFRQ' + 0.1$
    140 GOTO, STEP 105$
B       BRANCH OBJECT FROM STEP 120$
    150 DISPLAY, C 'REACHED 100.0KHZ WITH INPUT OF',
        'VLTFRQ', C' VOLTS'$
    160 FINISH$
    170 TERMINATE$
```

Figure 4

compiler (resident at PRD's Syosset, New York, facility in an IBM 1108 computer) by means of terminals located in the various VAST users' facilities. See Figure 5. After compilation a program tape and listings are returned to the user via the terminal. The control data on the tape is compressed. When the program is executed, this data is expanded in the system computer by means of the operating system, another computer routine. The operating system also supervises and controls the basic operation of the computer and peripheral devices. The operating system is stored on a tape and is read into the computer by means of a second Magnetic Tape Transport Unit.

The system block diagram, Figure 6, indicates the functional interrelationships of the system elements. MTTU #2 contains the tape upon which the operating system is stored. The operating system is loaded into the computer, a ruggedized Varian R-622/i. MTTU #1 contains the tape upon which the test program is resident. The Data Transfer Unit (DTU) is the primary man/machine interface. It contains a cathode-ray tube (CRT) display, status indicators, keyboard, and control switches. In order to execute the desired test program, the operator enters a message which identifies the desired program via the keyboard to the computer. At the beginning of each tape reel there is a listing of programs on the reel. The computer compares the requested program with this list, and if the requested program is not on the list, the operator is notified by a CRT message that the program cannot be found. If the program is on the tape, it is loaded. The beginning of the program identifies the building blocks required. These building blocks are placed in a full-power mode ready for operation. To enhance reliability, building blocks are normally in a warmup mode which energizes those circuits requiring more than fifteen seconds warmup time.

After the program is loaded into the computer the operator starts the test program execution. Digital commands are now transmitted from the computer on a ready/resume basis; i.e., a command word output from the computer is sustained until the next command word is requested. The Data Transfer Unit serves as a buffer between the computer and the building blocks, which have a standard logic interface. This buffering action of the DTU allows the building blocks to be independent of the system computer selected; changes to the computer can be made compatible with building block hardware by changes to the DTU.

All building blocks are connected to a common control trunk cable. When the DTU transmits a control word it appears at all building block control inputs. To differentiate between the commands, an addressing system is used. Each building block has a different address. When the test program utilizes a particular building block, the computer transmits an address command. All building blocks receive the address command; however, the effect upon the building blocks varies. The building block with that address transmitted places itself into a state to respond to subsequent commands. The other building blocks isolate themselves from the control trunk cables.

Upon enabling itself, the addressed building block transmits a verification signal back to the DTU and computer and the next command word is raised. Once again the word is held on the lines until the verification signal, indicating a response to the command by the building block, is generated by the building block and transmitted back to the DTU and computer.

Other signals are also transmitted back to the DTU and computer from the building blocks. Measurement blocks transmit serial data back for comparison in the computer and display upon the DTU cathode-ray tube. A printout of the display can be obtained from the input/output unit. Fault monitors within the building blocks check certain critical functions. Upon failure, the building block generates a fault signal which is transmitted to the DTU and computer. This activates a subroutine in the operating system which identifies the faulty building block, the nature of the fault, and displays it upon the DTU.

One of the original goals for VAST was the reduction in the number of skilled personnel required in the maintenance shop. For that reason the maintenance of VAST itself has been emphasized in the design implementation. The first level of testing is the ready/resume control technique itself. In most building blocks the verification signal is dependent not only upon receipt and recognition of the control word but also upon its execution. For example, the power supplies will not generate a verification signal or close their outputs until the voltage achieves the programmed value; signal generators will not verify until the frequency control loops achieve a stable, locked condition. Without verification, the test program cannot proceed. This condition is indicated on the DTU and the faulty building block is identified by means of an indicator on its front panel.

The fault lines previously mentioned provide a second means by which faults are detected and localized. While the verification signal detects faults only at the time the building block is being commanded, the fault lines continuously monitor and indicate building block performance.

In addition to these hardware features, there is a comprehensive software maintenance package. At the overall

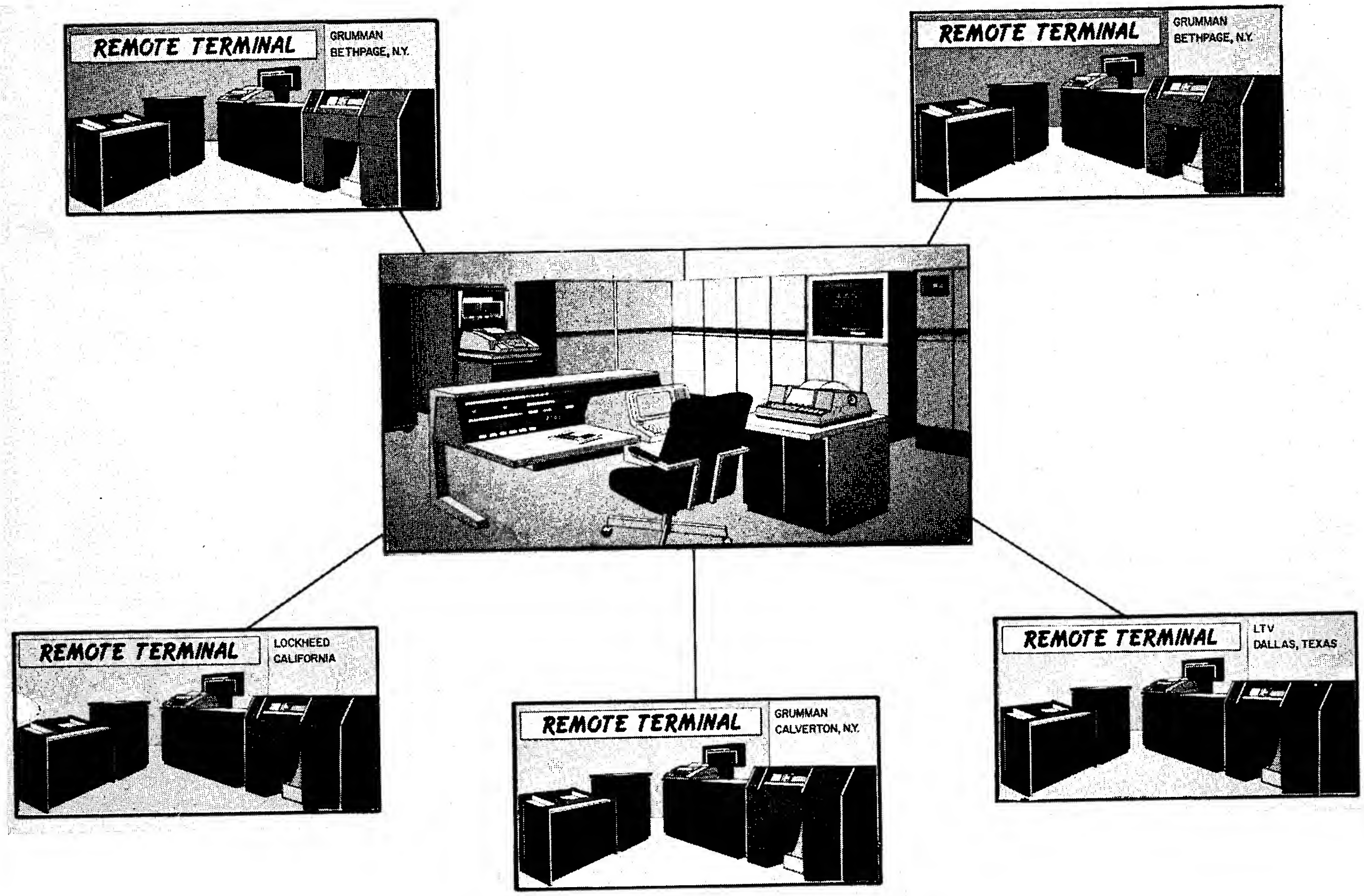


Figure 5

system level there is a self-check program. This program automatically checks out the operational suitability of VAST and isolates any failures to the next lower system element, e.g., a building block is isolated to an assembly by means of a self-test program which also will isolate failures to components within the faulty assembly. This series of programs also contains the calibration, adjustment and alignment procedures required to support the system. By means of these test programs VAST is virtually self-supporting.

VAST has achieved its original goals. It is currently being used by Navy personnel on board the USS Kitty Hawk supporting various A-7E electronics. In addition, VAST systems are located at Grumman Aerospace Corporation, Lockheed Aircraft Corporation and LTV where they are being used to develop test programs for the F-14A, E-2C, and S-3A, respectively. This broad applicability is indicative of the success achieved in developing a standardized test vehicle. Navy personnel were successful in the use and maintenance of the VAST System utilizing the available maintenance tools.

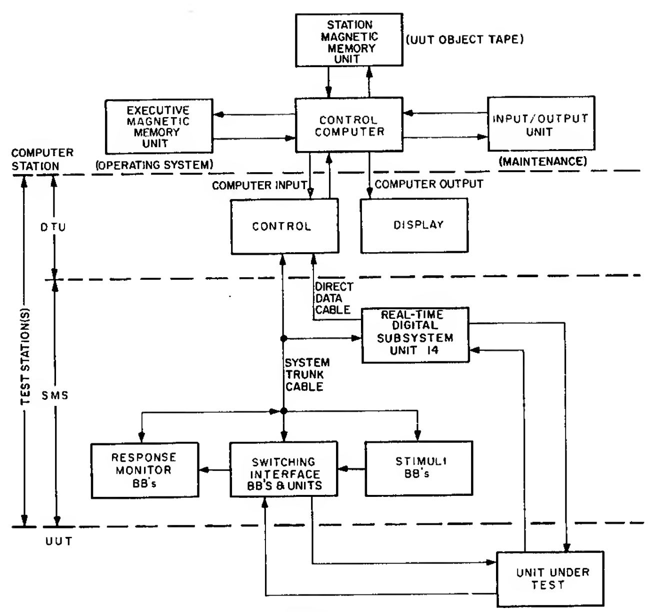


Figure 6

VEHICLE AVERAGE USEFUL LIFE STUDY

Raymond Bell
Robert E. Mioduski
U. S. Army Materiel Systems Analysis Agency
Aberdeen Proving Ground, Maryland

Introduction

An important consideration in the management of a fleet of vehicles, military or commercial, is knowledge of the useful life of the vehicles and whether or not it is economical to extend a vehicle's life by subjecting the vehicle to a costly major overhaul.

The Department of the Army in a move to reassess the useful life of its tactical wheeled vehicles requested the Army Materiel Command (AMC) to conduct a Vehicle Average Useful Life Study which would have the following primary objectives:

- Determine the age at which it becomes economical to replace each of the four major payload tactical wheeled vehicles (1/4, 3/4-1 1/4, 2 1/2 and 5-ton vehicles).

- Determine the economics of overhauling each of these wheeled vehicles and the remaining vehicle life after overhaul.

This paper will concern itself with the vehicle average useful life study conducted for the 2 1/2 ton vehicle. The results of this study, as indicated in this paper, should not be considered at this time as the official U. S. Army position on this subject.

Data Source

The data source being utilized in this study consists of two separate Army data collection systems: (1) The Army Equipment Record System (TAERS) and (2) Sample Data Collection Program. The TAERS data collection system for vehicles was instituted by the Army in 1963 and was designed to collect detailed maintenance information on all vehicles in the U. S. Army fleet. This data collection system, however, was terminated in December 1969. The Sample Data Collection Program for vehicles was initiated in 1972 and was also designed to collect detailed maintenance data, however, only for a sample portion of the wheeled vehicle fleet. The Sample Data Collection Program also differs from the TAERS system in that U. S. Army Tank-Automotive Command (TACOM) technical representatives are in the field in order to insure more complete and accurate reporting of data than occurred with the TAERS data collection system.

The TAERS data, which is the currently existing field data base can only be utilized for objective one (vehicle useful life) listed above as no substantial quantity of data exists in TAERS for overhauled 2 1/2-ton vehicles (M35A2 model). Data on overhauled 2 1/2-ton trucks will be collected in the Sample Data Collection Program and thus objective two (economics of overhaul) will be ascertained when this data is available.

Of critical concern in the use of TAERS data for analysis purposes, is the fact that many of the vehicle histories contained in the data bank are incomplete. This data omission problem is readily evident when vehicle histories are observed which shows, for example, for a truck produced in late 1965 only one maintenance action reported in the time frame 1966 thru 1969. As regularly scheduled maintenance actions (at least semi-annually) must have occurred with this truck during the '66 to '69 interval which should have been reported (scheduled as well as unscheduled maintenance actions are supposed to have been reported in the TAERS system) this truck obviously has incomplete data. Thus, in the use of TAERS, it is important that periods of incomplete vehicle histories be eliminated from consideration.

The method used by AMSAA to distinguish complete from incomplete periods of vehicle histories involved the TAERS quarterly reporting system. Under TAERS, a quarterly report of any maintenance actions (scheduled or unscheduled) occurring within the quarter was to be reported. Based on this requirement, selection of trucks for inclusion in the study had to meet the criterion that there were at least four quarterly reports in a row (one year of continuous data) in the truck history. This criterion although eliminating from consideration such vehicles as the one with one maintenance action in four years as well as vehicles with only intermittent reporting did not entirely resolve the data omission problem. Although the vehicles selected by this criterion had at least one year of continuous data, it doesn't necessarily imply that the vehicle's entire history was complete. For example, a vehicle produced in December 1965 may show TAERS reports in all four quarters in 1966 and the first three quarters of 1967 and subsequent to this period reports are indicated only for the third quarter of 1968 and the first and third quarter of 1969. Thus, after the third quarter of 1967 reporting became intermittent. The mileage noted on the vehicle during the first report in 1966 was, say 312 miles, with the mileage in the third quarter of 1967 being noted as 8,465 miles and the final mileage of 14,325 being noted by the report in the third quarter of 1969. If the missing quarters in 1968 and 1969 were ignored this vehicle history would assume to be complete through 14,325 miles. However, this may not be the case as maintenance actions may have occurred in the missing quarters of 1968 and 1969. Thus, for this study, that part of the history that provided only continuous reporting was used. In the above example, the vehicle's history only from 312 to 8,465 miles would be used.

Vehicle Sample

The data used in this study was obtained from TAERS reporting on 2,291 M35A2 2 1/2-ton Cargo trucks, 415 (18% of the total) were driven in Europe, 1575 (69%) were driven in the continental United States (CONUS) and 301 (13%) were driven in other parts of the world, primarily in the pacific area. The 415 European driven trucks covered 2.3 million miles, the 1575 CONUS driven trucks were driven 6.5 million miles and the 301 other trucks were driven 5.3 million miles for a grand total of 14.1 million miles for the 2291 trucks. The maximum mileage for an individual truck that was used in this study was 40,000 miles.

Useful Life Assessment Methodology

The useful life of the M35A2 2 1/2-ton Cargo Truck will be assessed by first determining the mileage at which the average system cost per mile (costs associated with the acquisition, shipping and maintenance of the truck) is minimized. This mileage at which the average system cost is minimized is called the economic life of the truck. In addition to determining the

economic life, an evaluation of the vehicle's Reliability, Availability and Maintainability performance characteristics over the economic life span is made to establish if the vehicle's useful life is less than the vehicle's economic life. For example, a truck at 30,000 miles may begin having frequent breakdowns due to a relatively inexpensive component failure. The effect of this type of breakdown may not be readily evident in a cost analysis alone but this breakdown may result in a substantial reduction in the vehicle's reliability. Of particular concern is the degradation of reliability below military requirements. If, however, the RAM parameters do not significantly degrade throughout the economic life of the truck, then the useful life could be equal to the economic life of the truck.

TAERS Data Analysis

In exercising the above methodology, the procedure employed was to analyze the maintenance costs (scheduled and unscheduled) to determine how the costs were changing as the vehicle increased in mileage. This procedure was also carried out for the analysis of the RAM characteristics.

The TAERS data utilized provided information on the maintenance actions (both scheduled and unscheduled) required for the vehicles as the vehicles increased in mileage. In particular, for each maintenance action, the following data were recorded: date action occurred, mileage at which action occurred, maintenance level (organization or support), man-hours required, failure detection code (i.e., whether the problem was detected in normal operation of the vehicle, during an inspection or is just a regularly scheduled maintenance action), remedial action taken (repaired, replaced, adjusted or is simply the result of normal services), part name and Federal Stock Number, and quantity of parts replaced.

The analysis of the data from a cost standpoint utilized the parts costs contained in the Army Master Data File. This cost information is in 1972 dollars and was supplied to AMSAA by TACOM. The labor rate used in this study was $4.39 an hour. It is noted that there were approximately 130,000 maintenance actions for the 2291 vehicle sample and about half of these were parts replacement. As noted earlier in this paper, data omission presented a serious problem in the analysis of TAERS. As a result of this problem many vehicle histories were incomplete. For example, the vehicle discussed earlier was considered to have a complete history only from 312 to 8,465 miles. Some vehicles had histories beginning at 10,000 miles and ending at 20,000 miles. In the costing of the maintenance actions by mileage, it was thus necessary to be aware of each vehicle's mileage interval. The costing procedure involved determining the total cost (parts and labor) experienced by the vehicles for each 100 mile interval. In this compilation, the vehicle with a history of 312 to 8,465 miles only contributed to the cost total beginning with the 300 to 400 mile interval and ending with the 8400 to 8500 mile interval. Thus, the sample size for each 100 mile interval is noted to vary. This procedure probably conservatively estimates the costs sustained as the vehicle which is noted to have its last maintenance action at 8,465 miles probably went many additional hundreds of miles without having to sustain any additional maintenance actions but in the procedure employed the vehicle was considered to contribute to the cost input up to 8500 miles only.

The analysis of the TAERS data from a RAM standpoint presented an additional problem. Normally in the analysis of data for the determination of reliability and availability estimates, failure data is required. However, from the TAERS data it is extremely difficult, if not impossible, to determine for all unscheduled maintenance actions which actions are reliability failures. As a result of this fact, an analysis of all unscheduled maintenance actions was undertaken. Specifically, the analysis consisted of three phases, all with the objective of determining how the vehicle's performance was changing as the vehicle increased in mileage: (1) Unscheduled Maintenance Action Analysis. The goal of this analysis was to determine the probability of completing a random 75 miles without an unscheduled maintenance for continually increasing mileages, (2) Inherent Readiness Analysis. The goal of this analysis was to determine the probability that the vehicle is not undergoing active repair due to an unscheduled maintenance action for continually increasing mileages and (3) Maintainability Analysis. This analysis consisted of computing for continually increasing mileages the maintenance support index (MSI), the average man-hours required per vehicle per 1000 miles of usage and the average man-hours required per maintenance action.

Cost Analysis

As noted earlier, the object of the cost analysis was to determine how the maintenance costs were varying as the truck mileage was increasing in order that the overall system costs could be minimized. Thus, all the maintenance actions occurring with the 2291 trucks in the study were costed (parts and labor) as a function of mileage. See Figure 1 for a summary of the costs as a function of mileage (in 1000 mile intervals) for mileages from 0 to 40,000 miles.

The methodology employed in the analysis of this data involved the application of weighted regression analysis techniques to the cumulative average maintenance cost results to obtain a continuous cumulative maintenance cost curve. The purpose of this determination was: (1) to obtain an initial cost, if any, for zero mileage, (2) to obtain a marginal or instantaneous cost curve (first derivative of the cumulative curve) and (3) to obtain an average maintenance and average system cost curve. From the instantaneous cost curve, the mileage at which the average system cost is at a minimum is determined. Further, 90% simultaneous confidence intervals on the mean cumulative cost, mean instantaneous maintenance cost and mean average system cost were computed.

In the analysis of the cumulative maintenance cost data, a third degree polynomial was found to best fit the data. Tests of significance of the coefficients indicated that the coefficients were highly significant (.01 level). The function determined was:

$$F(x) = 40.80 + 156.90X - 3.37X^2 + .0538X^3$$

where $F(x)$ = cumulative maintenance cost and

x = truck mileage (1000's of miles).

A plot of this equation with 90% simultaneous confidence intervals on the mean cumulative maintenance cost is shown on Figure 2. It is noted that the average cumulative maintenance cost for this truck through 40,000 miles of operation is $4,400.

The derivative of the cubic cumulative cost function which yields the instantaneous maintenance cost (or rate of change of the cumulative maintenance costs) is the following:

$$f(x) = 156.9 - 6.74x + .161x^2$$

where $f(x)$ = instantaneous maintenance cost and

x = truck mileage (1000's of miles).

Shown on Figure 3 is a plot of the instantaneous maintenance cost and the average maintenance cost as a function of mileage. The instantaneous maintenance cost is noted to decrease from 15.7¢ per mile when the truck is new to 8.6¢ per mile at 21,000 miles (the mileage at which the instantaneous maintenance cost is at a minimum) and to increase to 14.6¢ per mile at 40,000 miles. The average maintenance cost was found to be at a minimum at 31,500 miles (10.6¢ per mile) and averaged 11.0¢ per mile over 40,000 miles.

As stated above, the primary objective of this cost analysis was to determine the mileage at which the overall system cost to the Army is at a minimum, i.e., the costs associated with procuring, shipping and maintaining the truck are minimized. Utilizing the instantaneous maintenance costs developed and the truck rollaway cost (includes acquisition costs, engineering and tooling costs, administrative costs and first destination charges) of $10,861 plus a second destination charge (based on that part of the fleet that will be shipped overseas) of $307 an average system cost as a function of mileage is determined. A plot of the average system cost as a function of mileage is shown on Figure 4. As noted on the figure, the average system costs through 40,000 miles is still declining indicating the economic life of the truck is beyond 40,000 miles; however, if the trend in the maintenance costs developed through 40,000 is considered to continue beyond this mileage, then the average system cost is found to be a minimum at 60,600 miles. Also shown on Figure 4 are 90% simultaneous confidence intervals on the mean instantaneous maintenance cost along with system cost curves associated with these intervals. Thus a 90% confidence interval for the minimum average system cost of the 2 1/2-ton M35A2 truck is determined when the truck mileage is between 55,800 and 67,400 miles.

Performance Analysis

Unscheduled Maintenance Action Analysis

As indicated earlier, in place of a reliability failure analysis, an analysis of all unscheduled maintenance actions would be carried out due to the difficulty in determining in many cases if an unscheduled maintenance action was in fact a reliability failure. In analyzing the unscheduled maintenance actions, a system Weibull failure rate function was applied, i.e.,

$$r(t) = \lambda\beta t^{\beta-1} \quad t>0,\ \lambda>0,\ \beta>0$$

where t = mileage on vehicle

λ = scale parameter

β = shape parameter

This function assumes that the vehicle failure rate immediately prior to an unscheduled maintenance action and the rate upon completion of the action are the same and independent of the type of action performed. Furthermore, the failure rate determined by this function is also independent of the number of unscheduled maintenance actions previously performed on the vehicle. This differs from the standard use of the Weibull failure distribution which would assume that after each action the vehicle would be "as good as new."

From this function, the probability that a vehicle with mileage t will complete an additional s miles without undergoing an unscheduled maintenance action as determined by a non-homogeneous Poisson process is

$$P(s/t) = e^{-\lambda(t+s)^{\beta} + \lambda t^{\beta}}$$

where $\lambda(t+s)^{\beta} - \lambda t^{\beta}$ is the expected number of unscheduled maintenance actions for the interval [t,t+s].

The results of this analysis are shown on Figure 5. Indicated on this figure is the instantaneous unscheduled maintenance action rate and the probability of completing 75 miles without an unscheduled maintenance action for each 5000 mile interval from 0 to 40,000 miles. A goodness of fit criteria indicated that the model shown above highly represented the data. As can be readily observed on Figure 5, there is essentially no change in these parameters as a vehicle is increasing in mileage through 40,000 miles of driving. The average probability of completing 75 miles without an unscheduled maintenance action over the 0-40,000 mile interval is .96.

Inherent Readiness Analysis

As with a reliability failure analysis, the determination of availability is normally based on failure data. For example, Inherent Availability (A_i) is normally defined as:

$$A_i = \frac{MTBF}{MTBF + MTTR}$$

where MTBF is the mean time between failures and MTTR is the mean time to repair.

As noted in previous sections of this paper, unscheduled maintenance actions rather than failure data is indicated. Further, the TAERS data supplied information on the mean man-hours to repair rather than the mean time to repair. The mean time to repair for a particular maintenance action could be less than the man-hours involved if two or more men worked on a job. To utilize this data, however, to obtain some estimate of an availability statistic, one can determine the probability of a truck not undergoing active repair due to any unscheduled maintenance action when called upon to operate at a random point in time (Inherent Readiness) and this is given by the following expression:

$$R_i = \frac{MTBUMA}{MTBUMA + MMHTR}$$

where MTBUMA is the mean time between unscheduled maintenance actions and MMHTR is the mean man-hours to repair.

The Inherent Readiness parameter may be considered a lower bound on an Inherent Availability estimate, i.e., if all the unscheduled maintenance actions were reliability failures and if no more than one man ever worked on a maintenance action which would make the mean man-hours to repair equivalent to the mean time to repair then the $R_i = A_i$.

The results of this analysis are shown on Figure 6. Indicated on this figure is the mean miles between unscheduled maintenance actions (MMBUMA) and the Inherent Readiness, R_i (probability of truck not undergoing active repair due to an unscheduled maintenance action when called upon to operate at a point in time) for 1000 mile intervals through 40,000 miles. As can be readily observed, there is essentially no change in

the Inherent Readiness parameter as the vehicle is increasing in mileage through 40,000 miles of driving. It is noted that over the 40,000 miles the MMBUMA and Inherent Readiness are 1778 miles and .96, respectively.

Maintainability Analysis

The object of this particular analysis was to determine if the amount of maintenance required per truck was changing as the truck increased in mileage. Shown on Figure 7 is a summary of the maintainability data obtained for the M35A2 2 1/2-ton Cargo truck. Of particular interest in this figure is the average man-hours required per truck per 1000 miles, the average man-hours required per maintenance action and the maintenance support index (number of maintenance man-hours required per hour of truck operation); all reported by 1000 mile intervals through 40,000 miles of operation.

As can be readily observed on Figure 7, the average man-hours required per truck and subsequently the maintenance support index was noted to be decreasing through the first 20,000 miles of driving and was essentially level during the next 20,000 miles of operation. The first 1000 miles particularly required the highest number of man-hours. The basic data revealed that this was primarily due to the large number of man-hours associated with processing-in new vehicles. Overall, over the first 40,000 miles, the average man-hours required per truck per 1000 miles of operation was 9.2 man-hours, the average man-hours per maintenance action was 1.75 and the average maintenance support index was .18.

Summary

Useful Life Assessment

Based on 40,000 miles of operation, the average system cost has not yet reached a minimum thus indicating that the vehicle economic life is beyond this mileage. By assuming, however, that the trend in maintenance cost reflected during the first 40,000 miles of operation will continue, then the average system cost is minimized at 60,600 miles with a 90% simultaneous confidence interval of from 55,800 to 67,400 miles. Further, since none of the performance parameters, at least during the first 40,000 miles of operation were degrading as the vehicle mileage was increasing, the economic life noted may be considered the trucks useful life. If it is desired to convert the mileage indications to years, the M35A2 2 1/2-ton Cargo truck may be considered to have a 15 year life (based on 4,000 miles a year usage) with a 90% simultaneous confidence interval of from 14 to 17 years.

Profile of An Average Truck (over 40,000 miles of usage)

The average truck during the initial 40,000 miles of usage will sustain a total maintenance cost (scheduled and unscheduled) of $4400 or an average maintenance cost of 11.0¢ per mile at 40,000 miles.

During the 40,000 miles of usage, the average truck will have 22.5 unscheduled maintenance actions (UMA) with the mean miles between UMA's of 1778 miles. When the truck is in a maintenance shop for a UMA, on the average 2.3 parts will be repaired, replaced, or adjusted. During the average UMA 1.75 man-hours will be utilized for each part repair and a total of 4.0 man-hours will be expended for all repairs during an average UMA.

For each 1000 miles of vehicle usage, an average of 9.2 man-hours of maintenance (scheduled and unscheduled) are required. Of the 9.2 man-hours, 2.3 man-hours are for unscheduled maintenance and 6.9 man-hours are for scheduled maintenance. For every hour of truck operation (assuming an average speed of 20 mph), the truck on the average required .18 man-hours of maintenance.

For the average truck (over 40,000 miles of usage), there is at least a .96 probability that the truck will not be undergoing active repair due to a UMA and a .96 probability that the truck will complete a random 75 miles without a UMA.

FIGURE I

COST DATA FOR THE M35A2
2 1/2 - TON CARGO TRUCK

MILEAGE INTERVAL (1000's)	AVERAGE NO. OF TRUCKS	NO. OF MAINT. ACTIONS (SCH. & UNSCH.)	NO. OF MAN-HRS	TOTAL LABOR COST (DOLLARS)	TOTAL PARTS COST (DOLLARS)	TOTAL COST (DOLLARS)	AVG-COST PER TRUCK PER 1000 MILES (DOLLARS)
0-1	1,322	27,267	43,712	191,894	80,142	272,036	206
1-2	1,448	16,641	28,065	123,204	86,534	209,738	145
2-3	1,271	12,801	21,484	94,316	80,032	174,348	137
3-4	1,090	10,157	18,093	79,430	54,801	134,231	123
4-5	943	7,486	13,914	61,081	63,970	125,051	133
5-6	824	6,369	10,962	48,123	52,420	100,544	122
6-7	701	5,344	9,512	41,760	39,179	80,938	115
7-8	594	4,420	8,172	35,874	41,111	76,985	129
8-9	519	3,215	6,025	26,452	22,880	49,331	95
9-10	466	2,944	5,730	25,155	32,403	57,558	123
10-11	433	2,440	4,973	21,832	19,069	40,900	94
11-12	406	2,149	4,218	18,517	34,035	52,552	129
12-13	374	1,662	3,362	14,758	18,228	32,987	88
13-14	354	1,515	2,774	12,179	22,985	35,164	99
14-15	327	1,385	2,758	12,109	12,800	24,909	76
15-16	309	1,328	2,254	9,893	21,263	31,156	101
16-17	299	1,146	2,411	10,583	14,303	24,887	83
17-18	281	1,153	2,366	10,386	28,405	38,792	138
18-19	265	936	1,637	7,186	10,562	17,748	67
19-20	250	940	1,746	7,667	14,058	21,724	87
20-21	229	907	1,258	5,521	10,115	15,636	68
21-22	199	618	992	4,353	4,415	8,768	44
22-23	176	595	992	4,355	9,241	13,596	77
23-24	163	585	1,210	5,311	11,351	16,662	102
24-25	139	462	730	3,205	8,110	11,315	81
25-26	121	351	608	2,669	3,161	5,829	48
26-27	107	407	694	3,046	8,343	11,390	106
27-28	89	243	327	1,434	2,203	3,637	40
28-29	80	270	573	2,514	15,151	17,664	220
29-30	71	250	440	1,930	6,082	8,012	112
30-31	64	246	389	1,708	8,981	10,668	167
31-32	56	244	404	1,774	4,127	5,901	105
32-33	48	167	235	1,032	7,443	8,475	174
33-34	41	178	364	1,598	4,225	5,823	141
34-35	34	113	182	799	700	1,499	43
35-36	30	69	197	864	422	1,286	43
36-37	26	111	130	571	1,311	1,882	72
37-38	20	80	115	503	627	1,130	54
38-39	13	108	138	604	657	1,261	92
39-40	11	18	23	101	74	176	16

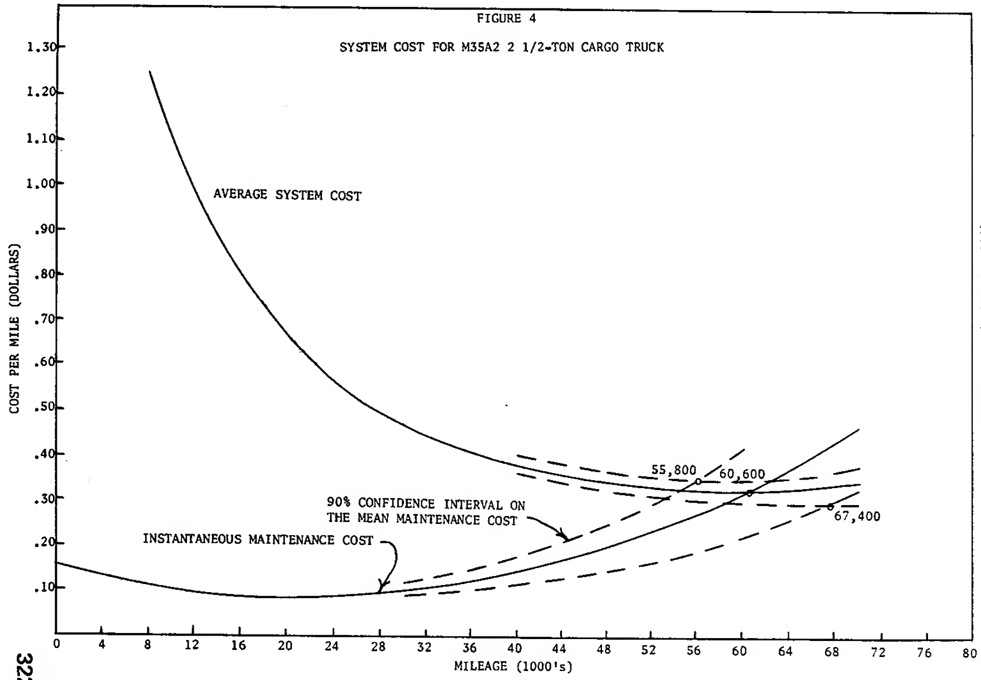


FIGURE 4

SYSTEM COST FOR M35A2 2 1/2-TON CARGO TRUCK

FIGURE 5

PROBABILITY OF COMPLETING 75 MILES WITHOUT AN UNSCHEDULED MAINTENANCE ACTION

MILEAGE	INSTANTANEOUS UNSCHEDULED MAINTENANCE ACTION RATE	PROBABILITY OF COMPLETING 75 MILES WITHOUT AN UNSCHEDULED MAINT. ACTION
0	-	.94
1,000	.00067	.95
5,000	.00060	.96
10,000	.00057	.96
15,000	.00056	.96
20,000	.00055	.96
25,000	.00054	.96
30,000	.00053	.96
35,000	.00053	.96
40,000	.00052	.96
AVERAGE	.00056	.96

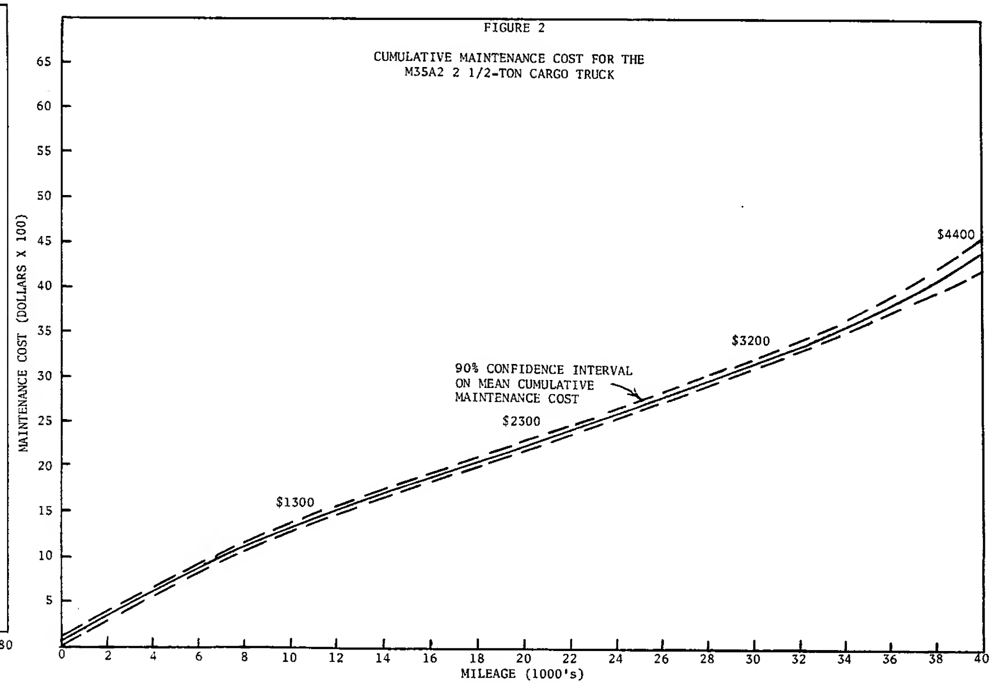


FIGURE 2

CUMULATIVE MAINTENANCE COST FOR THE M35A2 2 1/2-TON CARGO TRUCK

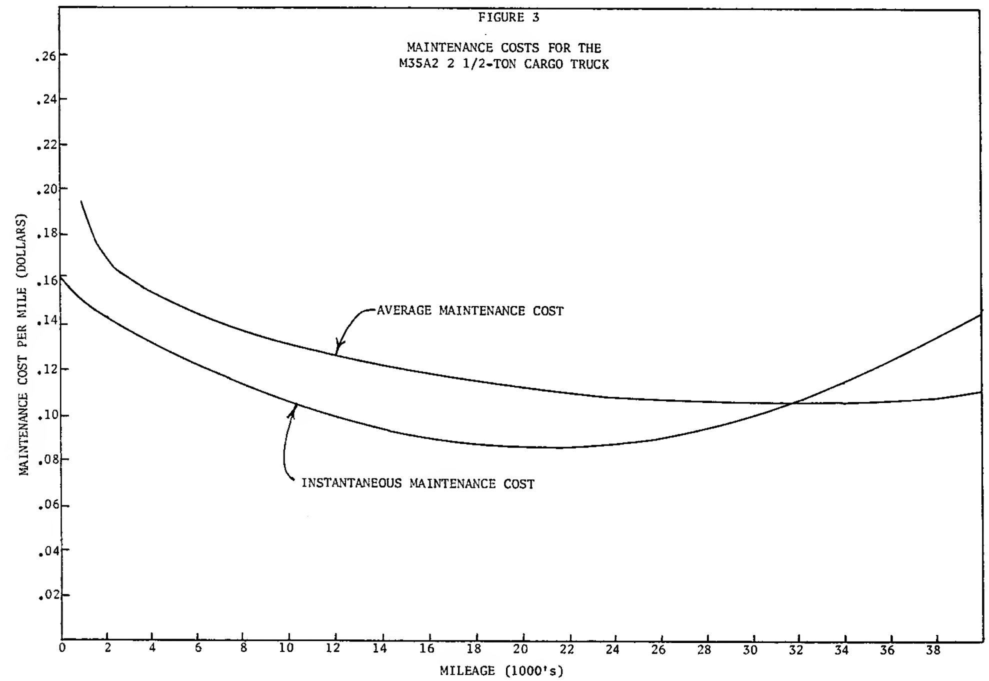


FIGURE 3

MAINTENANCE COSTS FOR THE M35A2 2 1/2-TON CARGO TRUCK

FIGURE 6

PROBABILITY OF TRUCK NOT UNDERGOING ACTIVE REPAIR DUE TO AN UNSCHEDULED MAINTENANCE ACTION AT ANY POINT IN TIME (INHERENT READINESS)

MILEAGE INTERVAL (1000'S)	MEAN MILES BETWEEN UNSCHEDULED MAINT. ACTIONS (MMBUMA)	INHERENT READINESS (R_I)	MILEAGE INTERVAL (1000'S)	MEAN MILES BETWEEN UNSCHEDULED MAINT. ACTIONS (MMBUMA)	INHERENT READINESS (R_I)
0-1	1385	.95	20-21	1823	.97
1-2	1525	.95	21-22	1829	.96
2-3	1580	.95	22-23	1835	.96
3-4	1617	.95	23-24	1840	.95
4-5	1645	.95	24-25	1845	.96
5-6	1668	.95	25-26	1850	.96
6-7	1686	.95	26-27	1855	.96
7-8	1703	.95	27-28	1860	.97
8-9	1717	.95	28-29	1864	.95
9-10	1730	.95	29-30	1869	.96
10-11	1742	.95	30-31	1873	.96
11-12	1753	.95	31-32	1877	.96
12-13	1763	.95	32-33	1881	.97
13-14	1772	.95	33-34	1885	.95
14-15	1781	.95	34-35	1889	.96
15-16	1789	.96	35-36	1892	.94
16-17	1796	.95	36-37	1892	.97
17-18	1804	.95	37-38	1899	.97
18-19	1810	.96	38-39	1903	.97
19-20	1817	.96	39-40	1906	.97

OVERALL MMBUMA = 1778 MILES

OVERALL R_I = .96

FIGURE 7

MAINTAINABILITY DATA FOR THE M35A2 2 1/2-TON CARGO TRUCK

MILEAGE INTERVAL (1000's)	AVERAGE NO. OF TRUCKS	NO. OF MAINT. ACTIONS (SCH. & UNSCH.)	NO. OF MAN-HOURS	AVERAGE MAN-HOURS PER TRUCK PER 1000 MILES	AVERAGE MAN-HOURS PER MAINT. ACTION	MAINT.* SUPPORT INDEX
0- 1	1,322	27,267	43,712	33.0	1.6	.66
1- 2	1,448	16,641	28,065	19.4	1.7	.39
2- 3	1,271	12,801	21,484	16.9	1.7	.34
3- 4	1,090	10,157	18,093	16.6	1.8	.33
4- 5	943	7,486	13,914	14.8	1.9	.30
5- 6	824	6,369	10,962	13.3	1.7	.27
6- 7	701	5,344	9,512	13.6	1.8	.27
7- 8	594	4,420	8,172	13.7	1.8	.27
8- 9	519	3,215	6,025	11.6	1.9	.23
9-10	466	2,944	5,730	12.3	1.9	.25
10-11	433	2,440	4,973	11.5	2.0	.23
11-12	406	2,149	4,218	10.4	2.0	.21
12-13	374	1,662	3,362	9.0	2.0	.18
13-14	354	1,515	2,774	7.8	1.8	.16
14-15	327	1,385	2,758	8.4	2.0	.17
15-16	309	1,328	2,254	7.3	1.7	.15
16-17	299	1,146	2,411	8.1	2.1	.16
17-18	281	1,153	2,366	8.4	2.1	.17
18-19	265	936	1,637	6.2	1.7	.12
19-20	250	940	1,746	7.0	1.9	.14
20-21	229	907	1,258	5.5	1.4	.11
21-22	199	618	992	5.0	1.6	.10
22-23	176	595	992	5.6	1.7	.11
23-24	163	585	1,210	7.4	2.1	.15
24-25	139	462	730	5.3	1.6	.10
25-26	121	351	608	5.0	1.7	.10
26-27	107	407	694	6.4	1.7	.13
27-28	89	243	327	3.6	1.3	.07
28-29	80	270	573	7.1	2.1	.14
29-30	71	250	440	6.1	1.8	.12
30-31	64	246	389	6.1	1.6	.12
31-32	56	244	404	7.2	1.7	.14
32-33	48	167	235	4.8	1.4	.10
33-34	41	178	364	8.8	2.0	.18
34-35	34	113	182	5.2	1.6	.10
35-36	30	69	197	6.5	2.9	.13
36-37	26	111	130	5.0	1.2	.10
37-38	20	80	115	5.5	1.4	.11
38-39	13	108	138	10.0	1.3	.20
39-40	11	18	23	2.1	1.3	.04

*INDICATES NUMBER OF MAINTENANCE MAN-HOURS REQUIRED PER HOUR OF TRUCK OPERATION (ASSUMING AN AVERAGE SPEED OF 20 MPH)

SUMMARY

1. AVERAGE MAN-HOURS REQUIRED PER TRUCK PER 1000 MILES OF USAGE = 9.2 (FOR BOTH SCHEDULED AND UNSCHEDULED MAINTENANCE).
2. AVERAGE MAN-HOURS REQUIRED PER MAINTENANCE ACTION = 1.75.
3. AVERAGE MAINTENANCE SUPPORT INDEX = .18.

RELIABILITY IMPROVEMENT OF LAND VEHICLES

Maxwell E. Westmoreland
U.S. Army Materiel Command
Washington, D.C.

This paper presents a profile of the current efforts of the Army Materiel Command, the principal developer of Army equipment, to improve the reliability of military designed land vehicles.

Introduction

The total annual maintenance support costs for the active operational U.S. Army fleet of combat and tactical vehicles has been estimated at $883 million. For the tactical fleet of just under 270,000 vehicles, maintenance support costs amount to $573 million, with 52 million man hours expended to do the work. The combat fleet of 21,000 vehicles costs $310 million to maintain with just under 8 million man hours expended.[1] As these costs reflect, the Army is incurring a significant cost burden in parts and labor to sustain its tactical and combat vehicles. If this burden for parts and labor alone could be reduced by just 5 per cent, the Army could avoid expenditure of almost $37 million annually.

The Army Materiel Command (AMC) has mounted an all-out attack on reducing this cost burden. The objective of this offensive is the reliability improvement of land vehicles to increase mission effectiveness and reduce the costs of maintenance support. The principle weapons being used in the attack are programs for: (1) assuring achievement of reliability in the design stage, (2) upgrading reliability through component development, and (3) reliability improvement of selected operational equipment. With this introduction, a discussion of each effort follows.

Achievement of Reliability in Design

Data presented during the Reliability, Availability, and Maintainability (RAM) panel discussion at the 1972 Army Science Conference indicated that one dollar spent on reliability during early development saved $100 retrofit costs.[2] Statements of this nature reinforce AMC's conviction that reliability can be achieved most economically during product design. The problem facing AMC is how to effectively do this. This problem is being tackled primarily through reliability growth management and in-depth design reviews.

Reliability Growth

In a typical development program, the objective of testing is to discover failure modes of the hardware design. When a failure mode is discovered, corrective action is taken to change the hardware design to eliminate this mode. The changed hardware is again tested to verify that the failure mode was in fact eliminated. The iterative process of design, test, redesign, test - and so on - has the effect of progressively increasing reliability. This progressive increasing of reliability in relation to development time has come to be commonly known as reliability growth.

The following examples are offered to illustrate some experience with land vehicles.

M60A2 Tank. Figure 1 depicts the reliability growth achieved on the M60A2 tank. The initial development prototype achieved 32 mean miles between failures, or MMBF. With these results, an intensive effort was begun to improve the tank MMBF to 110 miles. This effort involved the redesign of the original prototype based on the results of in-depth reliability analyses. The redesigned tank subsequently demonstrated a MMBF of 122 miles in engineering and service testing. Based on these tests, further modifications are being made; and the initial production is predicted to reach a MMBF of 139 miles.

Heavy Equipment Transporter (HET). Growth predictions were made in 1968 for the HET mean miles between maintenance actions, or MMBMA. The design requirement of 128 MMBMA was projected to be achieved upon completion of the check test in 1972. Based on changes made to correct faults identified in engineering and service tests, the 128 MMBMA was achieved in the check test as projected. (Figure 2).

M151 1/4 Ton Utility Truck, or Jeep. Reliability growth has resulted from improvements applied to each successive production model. The fourth production model is under procurement. The first, second, and third models achieved 1023; 2,444; and 2,980 mean miles between failure (MMBF) respectively. The fourth production model is predicted to achieve a MMBF of 3930 miles.

It is axiomatic that reliability growth will result in any development program due to the iterative testing and correction of design flaws. The critical factor, however, is the rate at which reliability growth occurs during the development time frame. Factors such as the number of items under test, the time allowed for testing and retesting, and the effectiveness of the corrective action process influence the growth rate. Experience has shown the consequences of failure to control the reliability growth rate to be:

1. User rejection of the hardware for failure to meet reliability requirements.

2. Excessive program costs attributable to schedule slippage and expensive product improvements.

AMC management has recognized that control of reliability growth is crucial to successful development of land vehicles. Today, development managers must execute programs for positive control of reliability growth. The key features of these programs are the determination of initial design reliability and the programming of predetermined manpower and funding resources to bring the design to the next predicted level of reliability. For an insight into these efforts, the Mechanized Infantry Combat Vehicle (MICV) program will serve as an example.

The MICV program is structured for the achievement of high reliability through a dedicated reliability growth program based on test and redesign between generations of development vehicles, independent design reviews, and an active component test program.

The MICV engineering development contractor will be required to develop curves for predicting and controlling reliability growth of selected components and subsystems similar to that shown in Figure 3. In Figure 3, the solid curve depicts the ideal reliability growth pattern for a particular component. The other

two curves are plotted from test failure data generated as the component and system development progresses. The lower plot, represented by circles, defines the trend of the cumulative mean miles between failure achieved. It includes all failures which have occurred to date. This plot should, with time, approach the ideal growth curve. If it does not, corrective action is indicated. The upper plot, represented by the deltas, defines the upper boundary for reliability. It includes only those failures for which design solutions are impractical or not contemplated. In effect, this plot eliminates all failures for which design corrections have been made and should approach the eventual production reliability growth.

Examples of component development and independent design reviews for the MICV program will be presented later during the discussion of component development and in-depth design reviews.

In-Depth Design Reviews

The concept of the in-depth design review is being emphasized in AMC as a means to assure achievement of reliability in development. In the application of this concept, AMC is structuring design reviews to make them formal, disciplined, and extensive evaluations to provide a timely mechanism for incorporating the composite experience of the organization into the design. In conducting these reviews, the team approach is used. The team is composed of the design engineer and carefully selected technical experts who have no direct responsibility for the design, but who possess unique knowledge and background to be able to constructively critique the design. For example, the AMC design review team for a tank program could be composed of the project manager, as the designer, and technical experts from the Weapons Command, Electronics Command, Munitions Command, Tank-Automotive Command, Army Materiel and Mechanics Research Center, and Army Materiel Systems Analysis Agency.

Reliability engineers will have a unique role in support of the indepth design review team. They will provide the team an extensive analysis of the design from a reliability view point. This analysis will address items such as reliability predictions, estimates of the design's reliability growth potential, failure modes and effects, environmental effects, and parts application studies.

The design review concept is being applied to the MICV program in the following fashion. An independent team of individuals experienced in reliability and design will perform integrated design reviews at critical milestones throughout the program. The purpose of these reviews is to determine if the hardware is achieving the degree of reliability growth required during development. The reviews are scheduled to provide for incorporation of team findings into the basic vehicle design prior to the final release of drawings for fabrication of both prototype and production hardware.

Upgrading Reliability Through Component Development

As mentioned above in the discussion on reliability growth, a sound component development program is essential to assure the timely achievement of reliability growth. The objective of AMC's component development effort is to determine technological limitations in componentry and to develop design approaches to minimize or eliminate the effect of such limitations.

Through analysis of experience data on land vehicles, AMC has found that poor reliability of these vehicles can be isolated to hardware components such as cooling systems, electrical components, fasteners, differentials, suspension components, brakes, and air cleaners. Causes of poor reliability in these components are generally found to be under design, poor integration, and poor production practices. The thrust of the component development effort for each of these areas is as follows:

Cooling Systems. An engineering design handbook on cooling systems is being written. In addition, cooling system components such as radiators, belts, clamps, and hoses are being tested to ultimately develop component selection criteria to assure the design application of hardware that will function in the military environment.

Electrical Components. To achieve system balance, electrical systems are being selected which meet the requirements of both the vehicle and the electrical system. Better ways to protect components from unusual loads and surges are being investigated. Starter protection devices are being developed to assure against destruction of starters from prolonged engagement with a running engine. Solid state ignition systems are under study to replace the coil and breaker approach. A standard 24-volt battery is being developed along with a device to control charging rate.

Fasteners. Shock and vibration environments are being studied to determine the best fastening procedures for use in components and vehicle systems.

Differentials. Approaches to the design of balanced drive lines are being studied. Methods of describing proven differential characteristics in drawings and specifications are being examined.

Suspension Components. The life of suspension systems is being increased through basic research in shock absorber technology, improving rubber wear pads on road wheels, and developing better specifications for the spring elements of suspension systems.

Brakes. Application of sealed disc brake technology to military vehicles is under investigation.

Air Cleaner. Efforts are continuing to increase air cleaner efficiency and service life.

Another type of component development effort is that done in support of specific vehicles during development. For example, in preparation for the current MICV program, six prototype MICV-65 vehicles were developed, fabricated, and subjected to military potential tests. Although further development and production of these vehicles were not undertaken, they have provided a foundation for the current program and have been utilized as test bed vehicles for further development, test, and evaluation of specific components and subsystems required for MICV. As a result of this effort, both commercial and military engines and transmissions have been tested and are available to meet power levels of the MICV. Also, four competing prototype stablization systems have been tested through second generation hardware with all systems meeting performance requirements. Thus, component hardware is available to meet MICV performance requirements; and, furthermore, these components are either off-the-shelf or well within the state of the art. The MICV program manager is therefore able to devote considerable effort which might otherwise be required for performance, to addressing the reliability growth that he seeks.

Reliability Improvement of Selected Operational Equipment

The discussion thus far has dealt with measures being taken to assure achievement of acceptable reliability in land vehicles before they are placed in the field. It is appropriate now to look at what is being done in AMC to improve the reliability of land vehicles already in the field inventory.

AMC has initiated a program for reliability improvement of selected equipment, popularly known as RISE.[3] The RISE program represents a sysematic assessment of fielded equipment to identify components and subsystems with less than desired reliability and a companion effort to engineer and apply cost effective modifications either on the production line or during repair and overhaul. RISE consists of four phases - identification, analysis, action, and verification.

Identification. First, a comprehensive assessment of equipment components is conducted to pinpoint those with high failure rates and excessive maintenance. The components are then arrayed from worst to best in terms of cost. Work is then begun on improving the worst ones first.

Analysis. The second step is the analysis phase. Careful engineering analysis establishes whether or not the component or part can be redesigned to increase its reliability. If redesign is feasible, the economic life cycle cost of the newly designed component is developed and compared with the cost of continuing to use the existing component. This is a detailed analysis which considers the cost to design, produce, apply, and stock the new component as well as the savings in cost resulting from improved system reliability and reduced maintenance as a consequence of the use of the new component.

Action. Here the reliability improvements are arrayed in order of merit savings-wise. Those with the most potential for cost savings are funded, and the improvements are made through either production line changes or through repair and overhaul in the field.

Verification. This phase closes the loop by monitoring the test and field performance of the improved hardware to determine if the improvement is meeting expectations. The information gained through verification also provides a baseline for establishing requirements for further development efforts on componentry.

The RISE program is working well for land vehicles. Some solid improvements have been made and others are underway. Here are some examples:

T130 Track. The M113 Armored Personnel Carrier Family has been using the T130 Track with a mean miles to replacement of 2000 miles. By changing to 4140 steel, the track is achieving 5000 mean miles to replacement with indicated savings of about $20 million.

M35A2 2 1/2 Ton Truck. The M35A2 2 1/2 Ton Truck has achieved 2600 mean miles between failure (MMBF) in production testing. The next production model of this vehicle will provide for a bootless front axle, sealed brakes, lifetime lubrication for joints and wear pads, improved seals, and improved trunnion axle bearings. These changes are expected to increase the MMBF to 3700 miles and to reduce scheduled and unscheduled maintenance by 11 and 34 manhours per vehicle, respectively, over the 20,000 mile life of the M35A2.

M114 Armored Command and Reconnaissance Carrier.

Historical experience on the M114 indicates that the mean miles between failures (MMBF) is 340 miles. By changing the engine, transmission, steer unit, and suspension system; the MMBF is expected to improve to 600 miles. These changes will also increase vehicle durability by 36 per cent.

AVDS-1790-2A Engine. This engine is the power plant for our M60 series battle tank. Historical experience on 4.6 million vehicle miles shows a mean miles between failures (MMBF) of 525 miles. An intensive effort is now underway to increase the engine MMBF to 1,379 miles by improving high mortality components. These improvements will increase the operational availability of the engine from 77 to 92 per cent and will extend the mean mileage to overhaul from 3180 to 3682 miles. Operational costs are expected to decrease by $34,000 over the 10 year life of each engine.

Conclusion

Within AMC, reliability achievement is considered the keystone for assuring the development of mission and cost effective equipment. Significant progress is being made in the application of the reliability discipline to this end. Noteworthy of these efforts are reliability growth management, indepth design reviews, component research and development, and reliability improvement of operational equipment.

References

1. US Army Tank-Automotive Command Report, Yearly Maintenance Support Costs for the US Army Vehicle Fleet, Revised May 1971.

2. Report of the Panel on Accelerated Development of Reliability to the Commanding General, AMC, A Study of the Army Materiel Command Reliability Program, June 28, 1972.

3. Army Materiel Command Regulation 702-15, Reliability, Availability, Maintainability Improvement of Selected Equipment (RISE), April 13, 1972.

M 60 A 2

RELIABILITY GROWTH

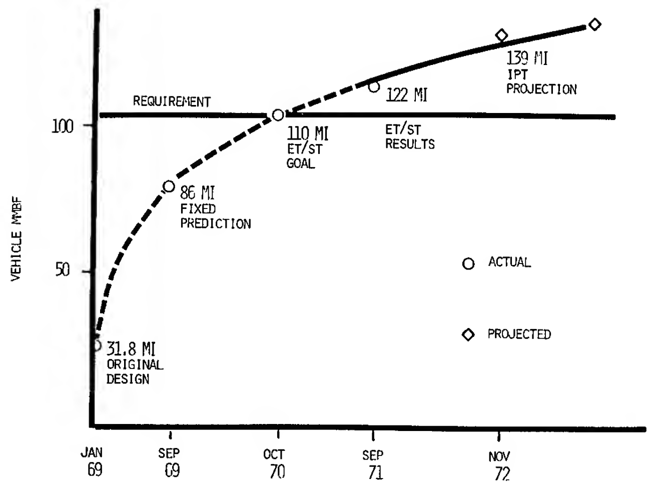
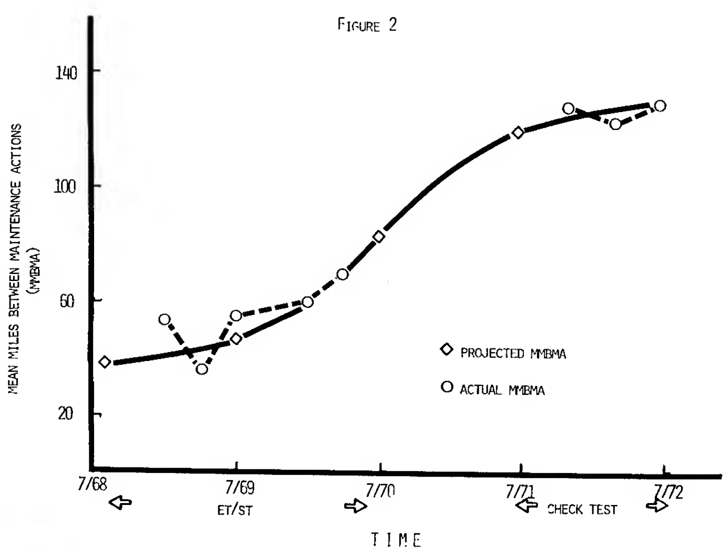
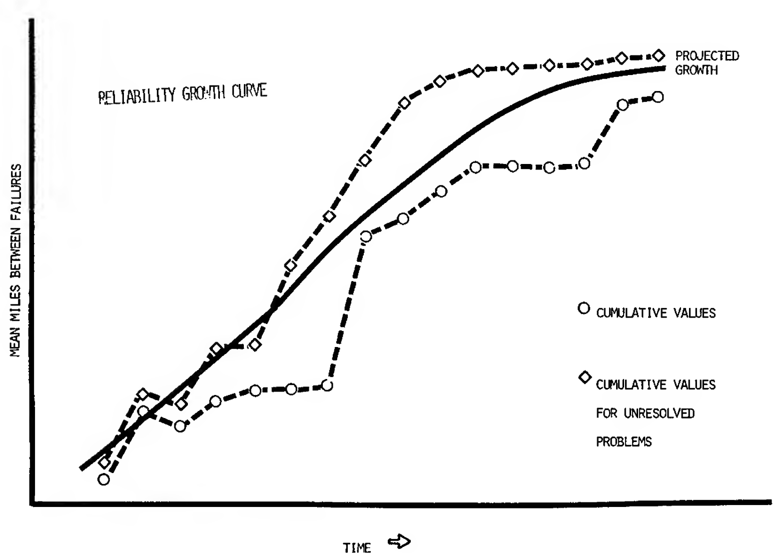


FIGURE 1

HEAVY EQUIPMENT TRANSPORTER





MAINTAINABILITY CONSTRUCTION MODEL ANALYSIS AND EVALUATION NETWORK

INDEX SERIAL NUMBER – 1088

William L. Andre
U. S. Army Weapons Command
Rock Island, Illinois

INTRODUCTION

The assessment of maintainability and maintenance characteristics of weapon systems is often a very challenging and enlightening experience, especially when the realization is made that reliability, performance levels, operating characteristics, hardware design, etc., all contribute to establishing those characteristics which define and constitute maintenance requirements. It was with this intention of establishing some useful relations and analytic procedures for measuring maintainability and maintenance parameters that the model herein presented was developed. The model was initially developed to perform a maintainability type analysis of the Army XM140 30MM gun system currently under development by the U. S. Army Weapons Command. The concepts identified and the parameters addressed will allow the model to be applied to a large number of systems. Application is primarily intended for evaluation of a system which requires component renewals during the course of its operating life.

The model is discussed in three sections:

A. Model Concepts

B. Network Analysis Diagram

C. Functional Block Procedures

The section on Model Concepts explains the rationale of the model approach taken and discusses model decision criteria. Subsequently the Network Analysis Diagram is presented. The network establishes the algorithm or procedure with its interrelationships for accomplishing the model intent. And finally a brief explanation of fundamental function blocks within the model is presented which illustrates mathematical computations used to determine variables identified in the network diagram.

MODEL CONCEPTS

The overall purpose and utility of the model is the determination of system maintenance costs and requirements given a system design. A selected design will have associated with it certain failure rates, failure modes, interdependencies of components, and repair requirements. These design characteristics are very influential in determining maintenance requirements. It is the task of translating these design characteristics into maintenance requirements of time and cost that the model attempts to achieve. Given a certain design, the remaining method of influencing maintenance requirements is the selection of the maintenance criteria or policy to be followed in supporting the system.

Although there are many variations of detailed policy that can be established, only two basic categories of maintenance exist. These are scheduled maintenance and unscheduled maintenance. The model considers these two categories as variables to be determined. For calculation purposes a maintenance point is considered a renewal point, and the word scheduled denotes maintenance disassociated with the occurrence of failure, whereas unscheduled denotes maintenance performed in response to or as a result of failures. As would be expected the relative costs of scheduled and unscheduled maintenance determine if and when scheduled maintenance is to be economically performed.

In order to establish cost values there are several factors that must be considered. For scheduled maintenance, it is manhours, tasks, parts requirements, and component costs which are considered as inputs for determining a cost value and time value associated with each major component of the system for scheduled events. The procedure used is not complicated by having to consider any kind of stochastic process since the selection of scheduled maintenance is not determined by failure rates, but instead determines to a large degree the observed failure rates of components. Failure rates which can be reduced by means of scheduled maintenance are considered as sustained failure rates, and the appropriate equations are used to calculate these rates as a function of the inherent design failure rate and the time to scheduled maintenance.

For unscheduled maintenance, system failure modes, related state probabilities, and component failure dependencies are considered in addition to those parameters utilized for scheduled maintenance event evaluation of time (i.e. MTTR mean repair time) and costs. Failure dependencies are sometimes considered as secondary failures. An example of what is meant here by failure dependencies can be illustrated by a control mechanism or safety device. If such an item fails it is often the case that the items or components being controlled or protected will also fail. Systems which contain components with a high degree of failure dependency between components are very good candidates for scheduled maintenance, especially if components tend to be expensive.

Due to the possibilities of secondary failures, and also the loss of any existing economies of scale, the cost of an unscheduled maintenance event is considered as being equal to or greater than the cost of accomplishing the same component renewal on a scheduled basis. The term economy of scale is used to signify cost savings that can occur strictly due to the performing of maintenance tasks on a scheduled versus unscheduled basis.

The second decision criterion of the model can now be addressed. It is to minimize total system downtime. The model performs an analysis based on time utilizing the same routine and procedures as used to perform a minimum cost analysis.

After minimization criteria are established for each system component, a total system allocation is performed to establish total costs for maintenance. To perform this allocation a linear programming routine is used which will consider variable coefficients of the objective function. The objective function is expressed in terms of system downtime for each category of maintenance considering dollar values for maintenance as the allocated variable to be determined. A set of constraint equations expressing the constraints placed on each component are used based upon the minimization conditions established. Dollar allocations are made until the minimum requirements of each system component are satisfied. The

allocation process can become an iterative process if the dollar value placed on downtime is considered variable with respect to the amount of funding considered available for maintenance.

Maintenance Decision Criteria

The decision to perform scheduled maintenance is established based upon whether or not it is economically feasible. To evaluate this condition, the average cost per unit of utility is considered. This unit may be time, cycles of use, etc. For uniformity of presentation the unit of utility is represented as time. The average cost C is calculated as:

$$C = \frac{C_u}{\mu_u} + \frac{C_s}{\mu_s} + C_o \tag{1}$$

where

C_u = cost of an unscheduled event

C_s = cost of a scheduled event

μ_u = mean recurrence time to an unscheduled event

μ_s = mean recurrence time to a scheduled event

C_o = a fixed cost value

The average cost expression is evaluated for each component identified in the system under consideration.

It is considered that the occurrence of either scheduled or unscheduled events constitute effective component renewal, and the expressions for the expected times to events can be calculated.

For scheduled maintenance events the mean time to scheduled maintenance is evaluated as:

$$\mu_s = \frac{\int_0^T R(t)dt}{R(T)} \tag{2}$$

and for unscheduled maintenance events (corresponding to failures) the mean time to occurrence is evaluated as:

$$\mu_u = \frac{\int_0^T R(t)dt}{1 - R(T)} \tag{3}$$

where T is the scheduled maintenance variable of the component age at which it is renewed if it does not fail before such time, and R(t) is the component inherent reliability function. The reliability function is primarily a function of design and represents that which is obtained say from performing a life test on a sample of N similar components, and does not represent the actual observed reliability of a maintained system obeying the relationships of equations (2) and (3) which apply whenever the two maintenance categories considered are sources of renewals.

The minimum cost criterion is based upon the above equations and identifies the value of T for which equation (1) will be a minimum. Substituting (2) and (3) into (1) and rearranging will yield:

$$\frac{C}{C_s} = \frac{X + R(T)(1-X)}{\int_0^T R(t)dt} \tag{4}$$

where $X = \frac{C_u}{C_s}$

Equation (4) becomes a minimum with respect to T when:

$$R'(T)\int_0^T R(t)dt = R(T)[\frac{X}{1-X} + R(T)] \tag{5}$$

The value of T which satisfies equation (5) identifies the scheduled maintenance point. A minimum cost condition will exist whenever $X > 1$. Calculating the cost ratio X for a component allows for determining whether or not component scheduled events should occur and equation (5) is useful in establishing when it should occur.

The parameters μ_u and μ_s have been identified in terms of the reliability function. Since the inherent mean life parameter is also a function of R(t) and is expressable in the form:

$$\mu = \int_0^\infty R(t)dt \tag{6}$$

It is clear then that μ_u and μ_s will be determined by T and μ.

Figure 1. shows graphically how the average cost of maintenance will vary as a function of T and cost ratio X for the normal failure distribution. The ordinate or y axis is expressed as a cost factor F where:

$$F = \frac{C\mu}{C_s}$$

with μ = inherent designed life
and the abscissa or x axis is the time T to scheduled maintenance expressed in terms of standard deviations from the mean.

A visual inspection of Figure 1 shows that even though a minimum cost point exists for all cost ratio values of X which are greater than one, a significant percentage reduction does not exist unless values of approximately 2 exist for X. This corresponds to a T value approximately equal to μ the component mean life. The model network presented in the next section uses as a lower limit the criterion that if $X = 1.87$ the minimum cost point is $T = \mu$. For this reason, values of T greater than μ should not be considered. T values for minimum costs will usually fall in the range:

$$\frac{\mu}{3} < T < \mu \tag{7}$$

Consider now the second model criterion of minimizing downtime. A minimum cost point and a minimum downtime point usually do not exist at the same point. However, the same rationale used for arriving at minimum cost criteria can be applied to establishing minimum downtime criteria. The cost terms for scheduled and unscheduled events identified in equation (1) can be replaced with downtime values without changing the logic of the equation. Therefore the expression for average hours of downtime for an hour of operating time will be:

$$t = \frac{t_u}{\mu_u} + \frac{t_s}{\mu_s} + t_0 \tag{8}$$

where t = downtime hrs/operating hour

t_u = time to perform an unscheduled event

t_s = time to perform a scheduled event

t_o = a constant time value

Since variations can exist between the amount of total maintenance, a system requires and the amount which contributes to downtime, a decision must be made when exercising the model to calculate component maintenance time in accordance with the system logistic support concept. If all maintenance is performed by the system user then all of the maintenance time will be considered as contributing to system downtime. If maintenance is performed primarily at higher levels such as direct support or depot, then only a percentage of total maintenance will contribute to system downtime.

Figure 1 can then be used to demonstrate for components with normal type failure functions the minimum downtime condition, where ordinate or y axis values will be:

$$\frac{t\mu}{t_s}$$

The abissa or x axis values will remain the same, and the X ratios will be time ratios of unscheduled to scheduled maintenance event times.

NETWORK ANALYSIS DIAGRAM

The time and cost considerations of the previous section provides the rationale utilized in establishing minimum conditions. Figure 2 presents the model network. Blocks 1 and 2 of Figure 2 represent the analysis of scheduled and unscheduled maintenance respectively which must be performed to properly evaluate the time and cost terms required.

Blocks 5 and 6 perform the function of establishing how inherent failure times are influenced by the operational use of the system. Block 5 particularly addresses the influence of operational effects on component failures. This is a function of design, and each system analyzed will exhibit its own characteristic equations. An example of operational effects is the analysis for block 5 that was performed for the XM140 gun system. Part of the effects analysis was establishing the effect of gun temperature levels on gun component failure rate. Five gun components were significantly affected by temperature. These items were barrel, muzzle brake, receiver, barrel cam assembly, and drum assembly. An analysis of the design and test data generated during gun development allowed the following relationships to be established:

$$\lambda_m = \lambda_0[1 + b(R_m - R_0)] \tag{9}$$

where R_O = reference component temperature level

λ_0 = reference component renewal rate

R_m = component temperature level for a specific mission profile

λ_m = component renewal rate for a specific mission profile

b = a constant, characteristic of the component design

Reference level values for temperature and renewal rate were established based on the average expected use of the system. The b parameter represents a rate of change with respect to temperature which is a constant for each component within expected operating ranges. The R_m term is calculated for each different mission profile of gun use as:

$$R_m = \frac{r}{N_m} L(t_m, R) \tag{10}$$

where t_m = mission time duration

r = firing rate

N_m = total rounds fired during mission

and $L(t_m, R)$ is the mission temperature - time integral integrated over the mission duration which excludes the non firing periods of the mission.

$$L(t_m,R) = \int_0^{t_m} R[\sum_{i=1}^{n/2} u(t-t_{2i}) - \sum_{i=0}^{n/2} u(t-t_{2i+1})]dt \tag{11}$$

where u - unit step function

n = number of firing bursts during mission

Equation (9) through (11) constitute the temperature effect relations used for one system and represent one approach suitable for establishing block 5 functions.

All of the tasks of blocks 1, 2, 5 and 6 are preliminary to performing the intent of the minimum time and cost equations identified in the previous section.

The combined functions of blocks 7, 8 and 9 perform the minimizing functions. The network of Figure 2

illustrates for these blocks the case for cost minimization even though downtime minimization is performed by these same function blocks.

The last function blocks executed are blocks 3, 4, and 10. These blocks perform the task of determining maintenance dollars allocated among system components. Funds are allocated based upon two criteria:

a) Meeting the constraints on maintenance time established by the selction of scheduled maintenance criteria.

b) The value placed on each dollar expended for each type of maintenance.

Block 3 performs allocations on a total system level and consists of a linear programming routine. A constraints equation for a total maintenance time for each system component is written which establishes criterion a). Criterion b) is accomplished by function block 4. This block determines the value of the coefficients used in the objective function being minimized by block 3. By considering the value of each dollar expended for maintenance to vary with respect to the amount of funds expended, a cost indifference curve effect is created where the value to the government in dollars for each hour of downtime decreases as the total number of downtime hours required increases. This consideration effects the coefficients of the objective function to be minimized in block 3, and requires iterative passes through the block 3 program until a convergent solution is reached. Once total system allocation is accomplished, block 10 performs allocations down to the component level.

FUNCTIONAL BLOCK PROCEDURES

Block 1

Scheduled event time and cost values t_{is} and C_{is} are calculated for each component by summing the task times, and costs required to effect the scheduled events. Inputs of manhour and parts requirements, and associated costs must be provided.

Constant average time and cost values are also determined if they exist for the system. These constants represent costs associated with daily use such as routine inspections, adjustments, calibration, lubrications, etc. which are continuing fixed values. This block reflects the independent variable characteristics of maintenance policy.

Block 2

This block considers system failure modes and their state probabilities of occurrence. Time and cost values are calculated for each failure mode considered. Failure modes are associated with each component prior to summing times and costs. In this manner interdependencies between components are addressed. The time and costs associated with each component can then be considered as independent times and cost in all subsequent calculations in the model. Block 2 can be used to calculate inherent component failure times or they can be supplied separately to the model program. The calculation of secondary failure costs and maintenance time is determined and added to scheduled time and costs in generating total unscheduled maintenance information. In evaluating block 2 for the XM140 gun system an existing model used by the Aeronutronic Division of Philco Ford which is useful for establishing maintenance task sequences and dependencies was used in conjunction with the unscheduled maintenance time and cost terms, and the state probabilities to establish total expected times and costs for each failure mode.

Block 3

The allocation process of this block seeks to minimize the linear equation:

$$Z = \sum_{j=1}^{k} A_j X_j \tag{12}$$

where j - describes the dype of maintenance

A_j = coefficient of variable (hours/$)

X_j = total funds allocated for maintenance ($)

The Z function is an expression of the total system downtime and is subject to a set of constraint equations which will reflect the minimum cost criteria or the minimum downtime criteria established for system components. The form of the constraint equation is:

$$\sum_{j=1}^{k} a_{ij} X_j > \text{Min}_i \tag{13}$$

The subscript i denotes the system component, and the coefficient a_{ij} is calculated as:

$$a_{ij} = \frac{\lambda_{ij}}{\sum_{i=1}^{n} \lambda_{ij}} \quad \frac{t_{ij}}{C_{ij}} \tag{14}$$

where λ = renewal rate

n = number of components

The value of the term Min.j is calculated as:

$$\text{Min}_i = \frac{Nw}{T_i - \dfrac{T_i^3}{6\mu_i^2}} \quad \frac{\lambda_{iu} t_{iu} + \lambda_{is} t_{is}}{\lambda_{iu} + \lambda_{is}} \tag{15}$$

with N = total number of systems maintained

w = time period being costed

T_i = scheduled maintenance criteria for ith component

t = component maintenance time

u = denotes unscheduled

s = denotes scheduled

The expression:

$$T_i - \frac{T_i^3}{6\mu_i^2}$$

is a measure of the expected time to component maintenance utilized for the analysis of the XM140 gun, and is equivalent to

$$\int_0^T R(t)dt$$

for a triangular type of distribution.

Block 4

The dynamic response function for calculating the block 5 objective function coefficients is:

$$A_j = \frac{X_j}{K_j \sum_{i=1}^{n} \lambda_{ij}} \sum_{i=1}^{n} \frac{\lambda_{ij} t_{ij}}{C_{ij}} \qquad (16)$$

The K_j term represents a dollar figure associated with the cost indifference concept previously mentioned.

Block 5

The equations and relationships used for determining operational effects depends on the system being evaluated.

Block 6

Mission transformation essentially performs the conversion of units of measure for what is defined for each mission. The model attempts to translate variables of miles, rounds fired, cycles of use, etc. into equivalent time units thus enabling outputs to be expressed as maintenance hours and costs per year.

Block 7

In calculating sustained time to failure and scheduled maintenance the equations in block 7 of Figure 2 were derived from a triangular failure distribution function and were found sufficiently accurate for also describing a normal failure function considering the inherent mean time to failure μ to be approximately equal to three standard deviations, and for weibull functions with shape parameters greater than 1.5.

Block 8

This evaluates the minimum time and cost equation of (1) and (8).

Block 9

This block selects the ratio of μ to T for each component based upon the time or cost ratio value input.

Block 10

The total system allocations of time and cost are broken out to the component level using the expressions:

a) total maintenance time for the ith component calculated as

$$a_{ij} X_j$$

b) and the corresponding component cost calculated as

$$\frac{\lambda_{ij}}{\sum_{i=1}^{n} \lambda_{ij}} X_j$$

REFERENCES

1. Andre, William L., "Analytic Maintainability Models Costed and Uncosted Which Consider Maintenance Ratios in Optimizing System Availability and Effectiveness", US Army Weapons Command MA-71-1, AD-726-904, June 1971.

2. Pieruschka, E., Principles of Reliability, Prentice-Hall, 1963.

3. Shooman, M. L., Probabilistic Reliability - An Engineering Approach, McGraw-Hill, 1968.

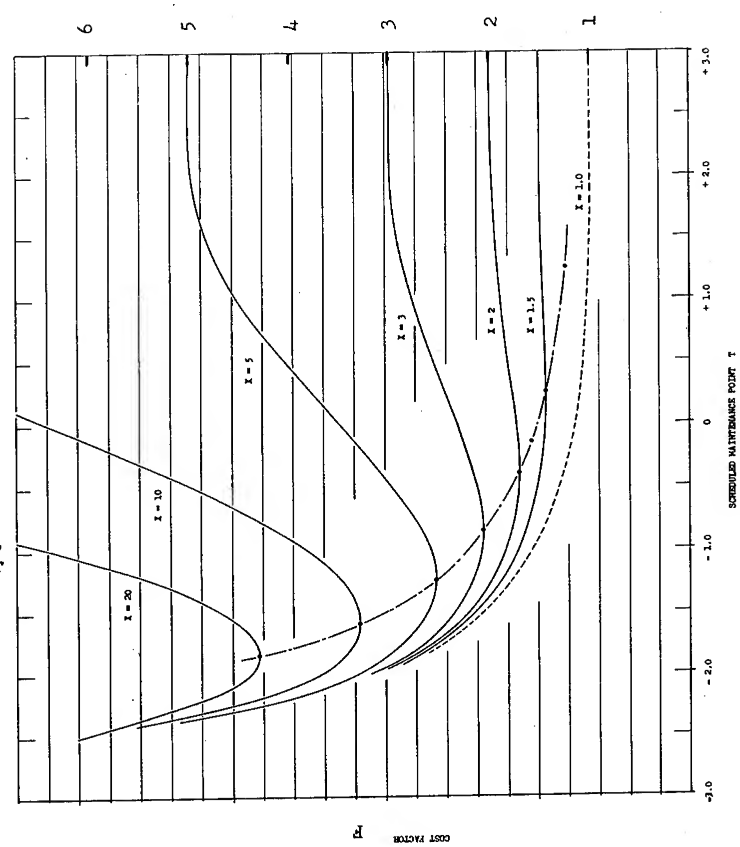
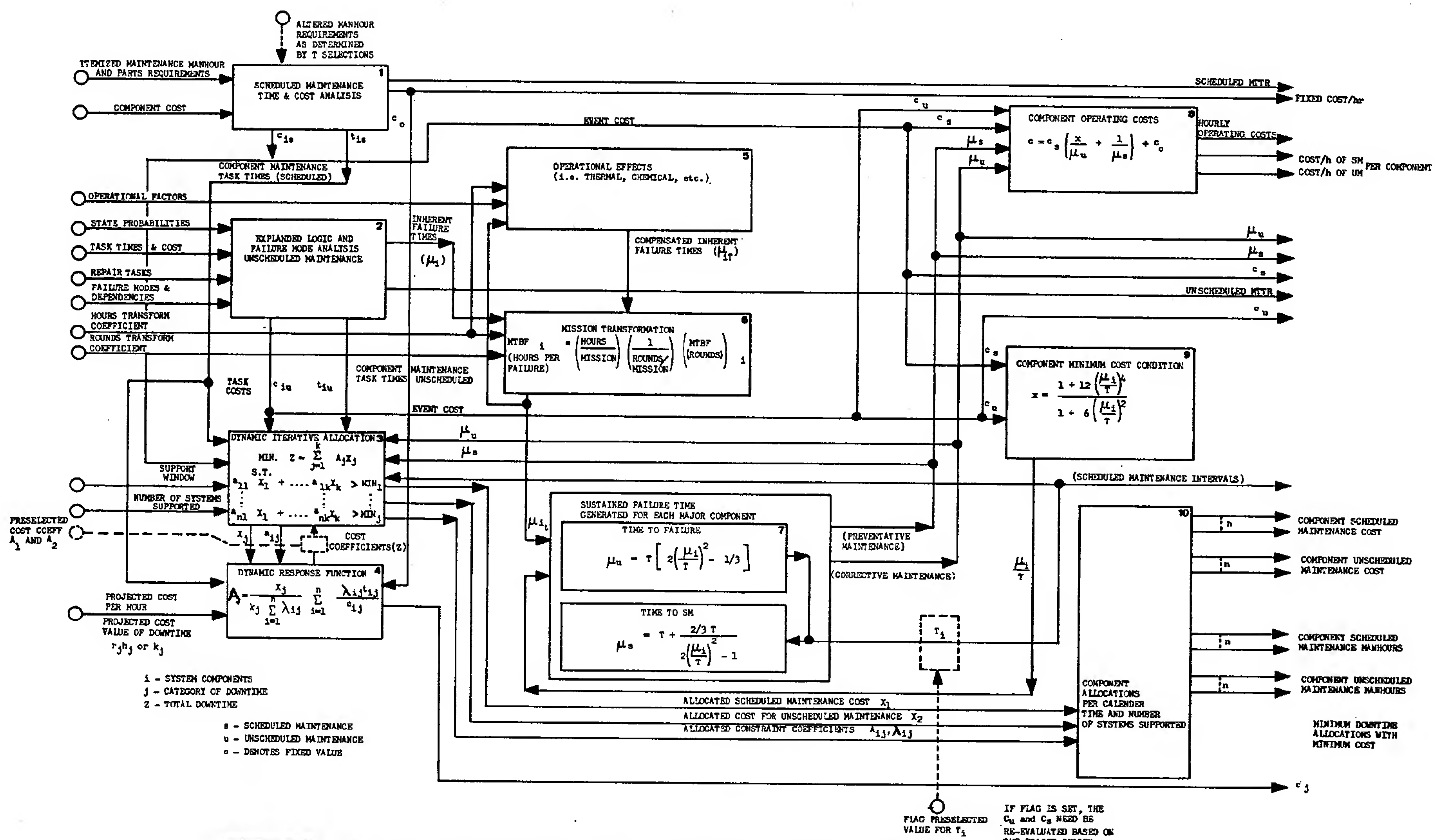
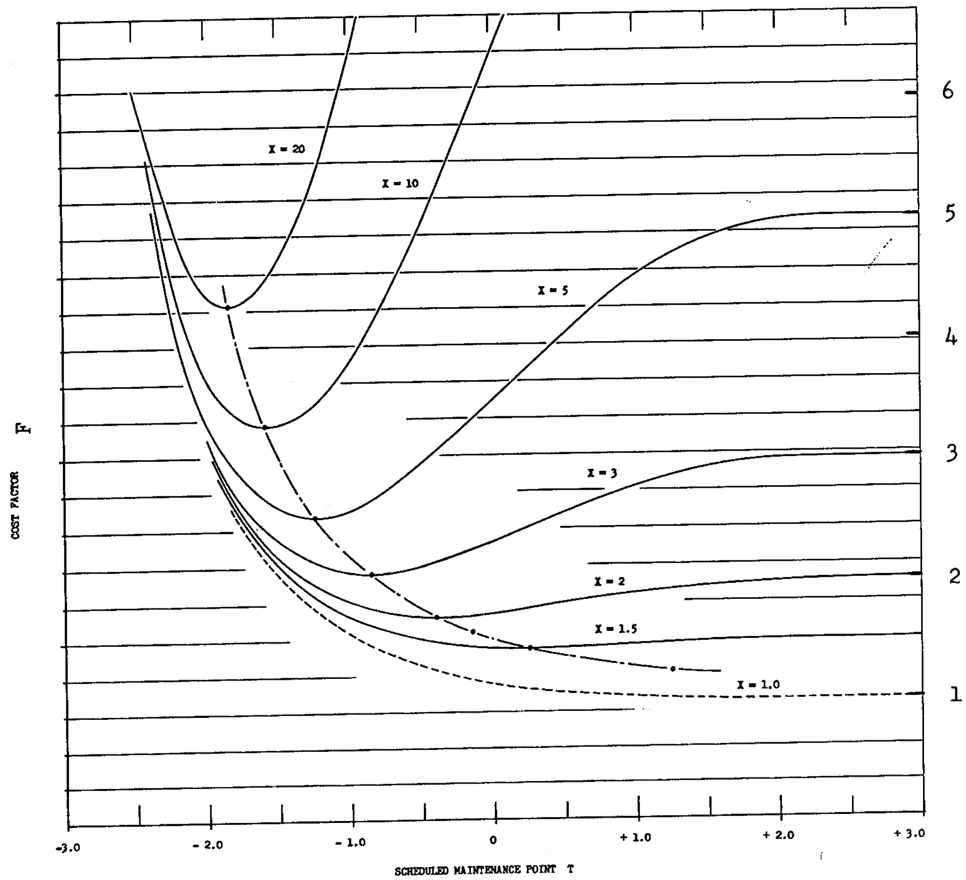


Figure 1



MAINTAINABILITY CONSTRUCTION MODEL ANALYSIS and EVALUATION NETWORK

Figure 2



COST FACTOR F
X = 20
X = 10
X = 5
X = 3
X = 2
X = 1.5
X = 1.0
6
5
4
3
2
1
-3.0
- 2.0
- 1.0
0
+ 1.0
+ 2.0
+ 3.0
SCHEDULED MAINTENANCE POINT T

Figure 1

HUMAN FACTORS AS IT AFFECTS RELIABILITY AND MAINTAINABILITY OF LAND VEHICLES

INDEX SERIAL NUMBER – 1089

John Erickson
U.S. Army Human Engineering Laboratory
Aberdeen Proving Ground, Maryland

When I hear the words maintainability and reliability, I visualize statements such as "the probability of this item performing successfully is 0.95" or "we estimate that the mean time between failures for this part is 500 hours or 2000 miles," and so forth.

Obviously, statements of this nature are predictions of what is expected to happen. The validity of the prediction is a function of the experience base upon which the prediction was made. The validity of the prediction is also a function of how the people who receive these items operate and maintain them. This is where human factors enters the maintainability and reliability picture.

We have observed that, under normal circumstances, the probability of an act being accomplished is a function of the ease with which it can be accomplished.

To a large extent this is a motivational factor. For example, the motivation of a vehicle operator has a direct effect on whether or not the oil level in the engine crankcase is checked in accordance with published requirements. At first glance, there doesn't seem to be too much you as a vehicle designer can do to motivate him. However, if he is inclined to check the oil level, there are several things you can do to maintain his motivation until he finishes this task. These include locating the dipstick in a position where he can gain access to it without having to take a grotesque or unsafe position or remove a number of bolts from an access plate or place his hand or arm in a position where it is subject to injury from either moving parts or hot objects, such as a manifold. On some of our vehicles, there is a requirement to drain contaminants from fuel and air lines. If these drain ports are improperly located in respect to access by the operator, he is less likely to, in fact, perform these tasks than if they were located so that he could access them from a normal and safe position. When the operator has to perform tasks of this nature in a driving rain, a snow storm, or when dressed in arctic clothing, he is less likely to perform the task than on a warm summer day.

It would appear that if the operator does not perform these tasks as prescribed, your maintainability and reliability predictions will go down the drain.

In regard to instrumentation in the vehicle cab, there are a few general comments I would like to make. In the first place, the only instruments that should be located in the vehicle cab, accessible to the vehicle operator, are those that are required for safe and effective operation of the vehicle. Second, these instruments should be located in a manner consistent with how the driver actually operates the vehicle. They should also be located in such a manner that they can be read by the operator with a minimum of interference with his primary visual task, keeping his eyes on the road. Unfortunately, we still see instances where an operator has to move his head and, in some instances, remove his hand from the steering wheel to see an instrument. Poorly located instruments will not be looked at. As a result, indications of marginal equipment operation may be overlooked with resulting larger maintenance problems occurring.

Since vehicles run just as well in the night as they do during the day, land vehicles are operated at night. Thus, if an instrument is required during the day, it, presumably, is required at night and thus comes the problem of instrument illumination. Unfortunately, there are too many examples of poorly illuminated instruments. The answer to this problem is not necessary one of increasing candlepower of the light bulb. The answer is more closely related to uniform illumination over the instrument face and from instrument to instrument. Another part of the answer is to provide large enough fiducial markings and numerals so that the operator can read the instrument from his normal eye position with low enough light intensity so that he is not temporarily blinded when he again looks at the road.

The final point I would like to make about instruments is this. Don't use complicated instruments. I believe it is safe to say that, in general, vehicle drivers are the low men on the pole in terms of training, time in grade, and education. The chances are pretty good that if a complex instrument is placed in the control cab, the operator will either make a mistake in reading it or he will ignore it. Therefore, why install it?

Warning plates and instruction plates are mounted in the driver's compartment of many vehicles. If it has been determined that information of this nature is required and has to be referred to while the vehicle is actually in operation, then--first, they should be designed and located in such a manner that the operator can read them while driving whether it is light or dark in the cab; and--second, the amount of information should be held to the minimum number of words and/or symbols that can be achieved without destroying their intelligibility.

In our more complex, crew-served vehicles, for example, a tank, there are a couple of other factors that come to mind. One is communications between crew members. If the communication system is not compatible with the noise environment inside the vehicle, the crew members will not be able to effectively communicate with each other. This may result in personnel being exposed to unnecessary safety hazards, making improper adjustments or misinterpreting instructions which result in damaged equipment. The other point is that as vehicles carry more complex equipment, it becomes more difficult for the crew members to isolate faults.

In those instances where on-board equipment can be designed for modular replacement, vehicle downtime can be reduced if the crew members can reliably identify the defective module and report this information to their support unit so that the correct parts and tools are brought to the vehicle on the first trip.

I think we have talked enough about the operating personnel, therefore, I'll mention a few points about maintenance. There is considerable human engineering criteria and data available that can and is being applied to maintenance. This includes improved

training techniques, improved job aids, task and skill analysis and design of tools, fixtures and instruments.

However, there is still the need to assure that when reliability predictions indicate that one component has a much higher failure rate than another, it should be located in the more accessible location for the obvious reason of reducing maintenance time. Bolts, screws and fasteners should normally be captive when they are difficult to reach and when they will create a problem if they are dropped. Quick-connect, quick-disconnect fasteners generally reduce maintenance time. Keying of electrical connectors, printed circuit boards, and various fluid connectors are necessary to reduce assembly errors, which, in turn, can become costly. The design of test instrumentation should be consistent with the accuracy required and the skill level of the intended user.

The design of fixtures and handling equipment should be consistent with the intended function they are to serve and they should also be compatible with the man who will use it.

One individual in the chain of events at a maintenance unit is the inspector. He is the individual who analyzes the vehicle, determines what is wrong and writes up the job order. He also inspects the vehicle after it has been repaired and then releases it either to supply or back to the user.

He can become a bottleneck and he can also cause unnecessary work to be performed if he misinterprets the cause of a failure.

As automatic fault analysis and detection equipment becomes available for automotive equipment, his ability to perform his task more rapidly and more accurately will be enhanced. However, considerable good human engineering will have to be applied to the development of this equipment to assure that he will have sufficient confidence in the equipment to use it. In addition, connection points should be located in the vehicle so that they can be accessed easily and rapidly. Operating the equipment should not be so complicated that he has to hold a book in his hand while using the equipment. If the equipment is properly designed, it will be accepted with gratitude; if, on the other hand, it is hard to use, or is itself unreliable, it will end up in the corner of the shop and only be brought out for open house and command inspections.

To illustrate the contribution human factors engineering can make to improving maintainability of military vehicles, I will spend a few minutes discussing our participation in the development of the XM-759, Marginal Terrain Vehicle (Figure 1).[1]

In August 1965, our Laboratory was requested to advise and assist the project engineering staff located at U.S. Army Tank-Automotive Center in the selection of the best of several concepts of a proposed marginal terrain vehicle (MTV).

Detailed reviews were made of the several concept drawings and, based on a human factors engineering (HFE) analysis of the mission and military characteristics as outlined in a U.S. Marine Corps (USMC) Specific Operational Requirements Document, the concept that would accommodate 14 fully-equipped troops within the cargo/troop transport area was recommended and, following an inprocess review on September 1965, approved by the USMC for development.

During the development of the first pilot vehicle, several visits were made to ATAC (now TACOM) to review the drawings and provide consultation to the project and design engineers on HFE practices and principles to assure that HFE was considered and incorporated in the vehicle design. As the program progressed from the drawing boards into the mockup state, review and evaluation of the full-scale model mockup was made to identify and correct HFE problems prior to fabrication of hardware and assembly of the pilot vehicle.

As a result of the periodic visits, the reviews and evaluations conducted, HFE inputs were incorporated that improved the crew workspace in the cab area so it would accommodate the full range of user personnel, 5th through 95th percentile; provided a better design of the brake and accelerator pedals; eliminated an interference problem in gear shift lever operation and reduced the possibility of inadvertent shifting into reverse gear by repositioning and redesign of the shift lever and shift quadrant. Also, improvement in the seat-control relationship and better forward and side vision for the vehicle operator were obtained.

The uniqueness of the vehicle design, plus the limitations and restrictions imposed by such design parameters as structural integrity, did create some HFE problems peculiar to this vehicle. Each of these problems was addressed and, although ideal solutions were not always achieved, such as the size of access openings to components in the engine compartment through the rear bulkhead, the best possible solution at that time was incorporated.

Some of the other areas addressed were cab area ingress and egress, cargo/troop area ingress, communications between personnel in troop area and vehicle operator, effective utilization of proposed armament for the vehicle, maintainability, stowage of on-equipment material (OEM) and life support items, utilization of kits, environmental factors such as vehicle noise, toxic fumes and climatic conditions, and safety.

One of the most difficult problems encountered on the early pilot vehicles was getting "maintainability" designed into the vehicle.

Limitations and restrictions imposed by the structural engineers, and the lack of appreciation by the design engineers for the need of "maintainability" as a design characteristics, resulted in very limited accessibility to components requiring daily servicing and checking operations and scheduled or unscheduled maintenance.

By continuously emphasizing the need for "maintainability" to be incorporated into the vehicle design, by periodic visits to contractor's plant to review and evaluate design drawings, discussions with the maintenance engineers, and, as the structural engineers became more knowledgeable in the application of the materials used on the MTV, considerable improvements in the design for maintainability were achieved. In comparison to the very limited access for maintainability on early pilot vehicles, the accessibility on pilot No. 7 is considered to be very good. Some of the salient access features on pilot No. 7 are: the engine compartment shroud, engine compartment cover, displaceable crew seats, displaceable steering control, large access on rear bulkhead, 18 inch diameter access in bottom of vehicle,

and modular cooling system with cooling fan and radiator mounted on the fender.

The engine compartment shroud, 60 inches long, 30 inches wide and 20 inches deep, is torsion-bar hinged on the left fender. The shroud is lined with acoustic material and has stowage and mounting for the radio and scrambler. The torsion-bar assist enables the 5th percentile to raise the shroud which locks automatically into position on the left fender (Fig. No. 2).

The engine compartment cover located between the shroud and crew seat is 54 inches long, by 15 inches wide and approximately 4 inches deep. It is hinged to the shroud and can be raised with the shroud to provide greater access to components under the shroud area or it can be raised independently along with the displaceable crew seat for access to components under the seats.

With the steering control displaced, the crew seats folded and displaced, the engine cover and shroud raised, there is an opening 44 inches wide by 74 inches long to provide an interference-free lift to facilitate power plant removal (Fig. No. 3).

On the first pilot vehicle, considerable difficulty was experienced in obtaining a 9 inch by 12 inch access opening in the rear engine bulkhead. However, by the time pilot vehicle No. 7 was fabricated, a 24 inch by 42 inch access opening was incorporated in the bulkhead, thus facilitating many maintenance tasks (Fig. No. 4).

Access for draining coolants and lubricants is provided by an 18 inch diameter opening under the engine compartment. Access to the fuel tank selector handle is through a quick release hinged cover at the top of the rear bulkhead.

Other maintenance features on pilot No. 7 included vehicle electrical wiring system centrally located with leads and connections marked and identified; quick disconnect electrical connectors on engine; ready accessibility of all filters (fuel, oil and air cleaner); batteries stowed in covered plastic boxes that are corrosive proof.

In August 1971, the HEL representative chaired a Maintainability Demonstration at the contractor's plant and directed the evaluators on the review and evaluation of the Draft Technical Manuals (DTM), Part, Maintenance Allocation Charts (P-MAC), On-Equipment-Material (OEM) lists, Special Tools list and several maintenance operations to determine the simplicity, clarity, completeness and adequacy of instructions, diagrams, photographs, illustrations, charts and lists. Discrepancies were noted and recommendations for revisions, changes and deletions were made.

Prime consideration for the stowage of items was adequacy of space for all items, ready accessibility under all operating conditions and vehicle stability. Stowage provisions for the tools, OEM, and life support items were considered early in the program and were kept current with requirement changes.

One of the requirements was to provide a means to transport several litters for a short distance. The original ATAC litter rack design was time consuming and difficult to install and required the removal and stowage of the troop seats. HEL addressed the problem by designing litter brackets that required only seconds to install, did not require the removal and stowage of troop seats and provided the capability of transporting five litters.

A heater-defroster kit for cold weather (-25°) operations and defrosting or defogging has been incorporated in the vehicle design.

A slave receptacle kit for auxillary electrical power for vehicle starting is provided and located for ready accessibility on the forward bulkhead.

In addition to providing crew protection from the climatic conditions in which the vehicle is required to operate and minimizing exposure to toxic fumes through cab design and discharge of exhaust outboard of the vehicle, HEL, at the request of AMC, conducted several studies[2,3], on the vehicle noise level to determine sources of excessive noise and methods of attenuating the noise to an acceptable level. Review of test data from various tests agencies and an early study on pilot No. 3 disclosed that the vehicle noise level exceeded HEL-STD-S-1-63B, "Maximum Noise Level for Army Materiel Command Equipment,"[4] by as much as 23 dB. Early attempts to reduce the noise level resulted in achieving a noise reduction that was considered acceptable but vehicle cooling requirements could not be met.

Complete redesign of the cooling system and continued work on noise attenuation resulted in meeting the vehicle noise level requirements and vehicle cooling requirements.

Safety considerations have been applied in the design of the MTV pilot vehicles from the beginning of the program and, as the program evolved and experience from field testing indicated methods or features that increased safety, these were incorporated in each of the subsequent pilot vehicles.

Some of the safety features on pilot No. 7 are: a safe method to perform daily draining of contaminants from the fuel system that eliminated a potential fire hazard; fuel tank fillers with spillway to direct overflow outboard and allow refueling in adverse terrain such as when operating in swamps, drain plugs incorporated in each fuel tank to eliminate the need of siphoning; handholds to assist in ingress and egress of cab and cargo area; foot steps in tailgate and nonskid material on walkways, steps and work areas; provisions for draining coolant and lubricants to outside of vehicle either directly or through use of drain hoses; covers and grills over hot surfaces to prevent burns and protect equipment, shrouding the cooling fan; exhaust fumes directed outboard away from personnel occupied area; use of wood for troop seats to reduce cold soak or radiated heat problems; fire extinguishers provided and located for ready accessibility; lifting eyes for heavy equipment; transmission neutral safety switch; and crew seat safety belts.

The U.S. Army Human Engineering Laboratory has actively participated in the XM759 vehicle program by incorporating proper and timely application of human factors engineering into the vehicle design from early concept drawings, through all development and testing phases and final technical data package. The end result is a vehicle designed insofar as possible so it can be effectively, efficiently and safely operated and maintained by the user personnel, within the limitations and restrictions imposed by the vehicle design and military characteristics.

In closing, I would like to say that reliability and maintainability can be enhanced when equipment is designed for the user to use. This requires an

appreciation of his motivational factors, the environment in which he lives, his physiological and psychological capabilities and his skill and training. A considerable amount of information is available today and the human factors community is continually developing additional information to aid the designer in these areas. However, a continuing effort on the part of human factors engineers supporting reliability and maintainability engineers will be required to assure that available information on human performance is applied to vehicle design to increase the probability that maintainability and reliability predictions are achieved when these vehicles are issued to the user.

References

1. Human Engineering Laboratory, Human Factors Engineering Participation on the Carrier, Cargo, 1 1/2 Ton, Soft Tire, Track, XM759 (MTV). HEL Letter Report No. 144, Human Engineering Laboratory, Aberdeen Proving Ground, Maryland, Aug 72.

2. Human Engineering Laboratory, Noise Reduction of the XM759 Cargo Carrier (Marginal Terrain Vehicle). HEL Technical Note 5-70, Human Engineering Laboratory, Aberdeen Proving Ground, Maryland, Jul 70.

3. Human Engineering Laboratory, Cab Noise in the XM759 Cargo Carrier (Pilot Vehicle No. 7). HEL Letter Report No. 139, Human Engineering Laboratory, Aberdeen Proving Ground, Maryland, Sep 71.

4. Human Engineering Laboratory, Maximum Noise Level for Army Materiel Command Equipment. HEL-STD-S-1-63B, Human Engineering Laboratory, Aberdeen Proving Ground, Maryland, Jun 65.

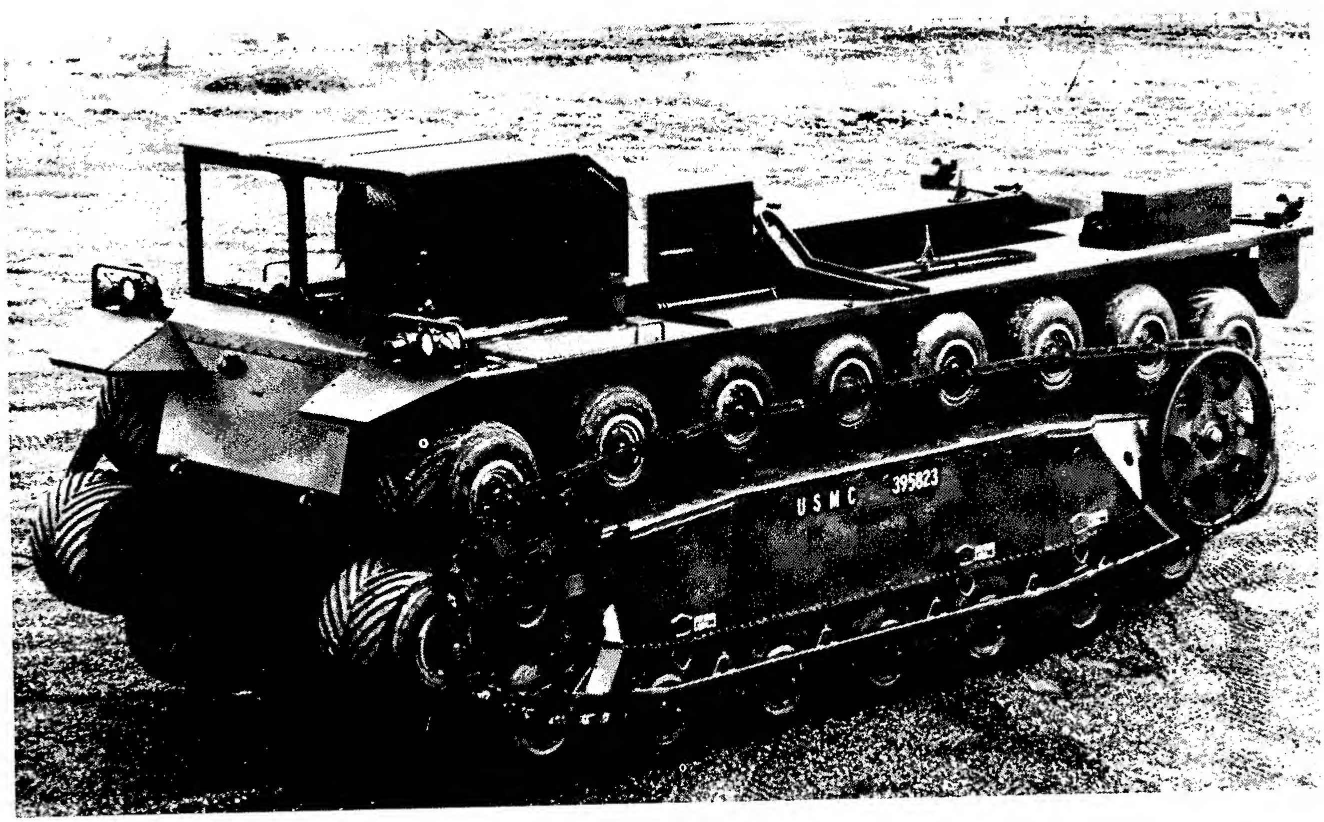


Fig 1. XM-759, Marginal Terrain Vehicle

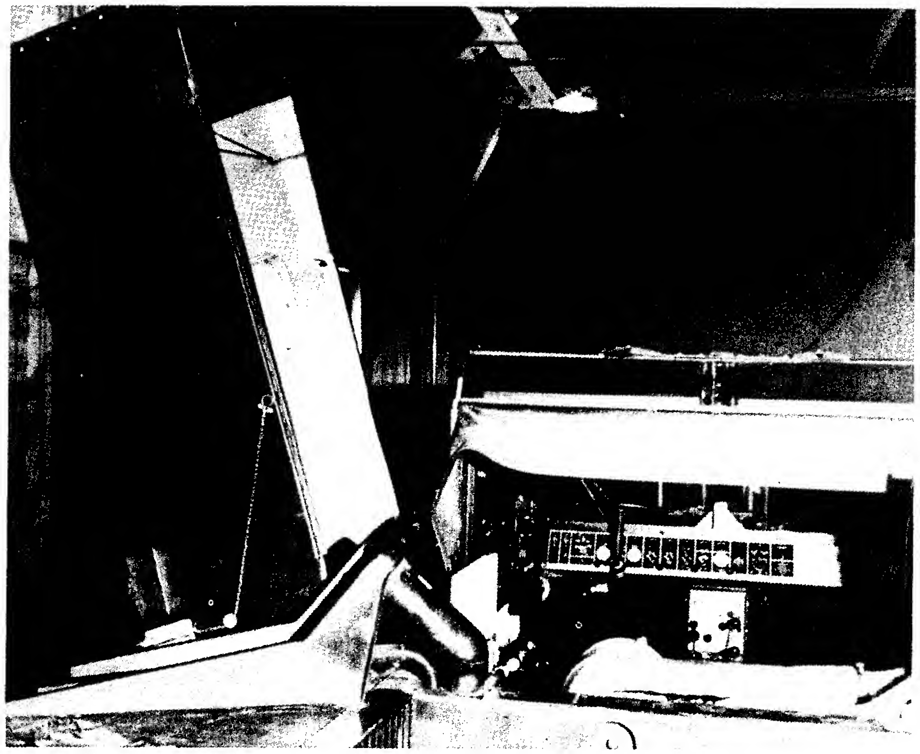


Fig 2. Engine Compartment Shroud, in open position

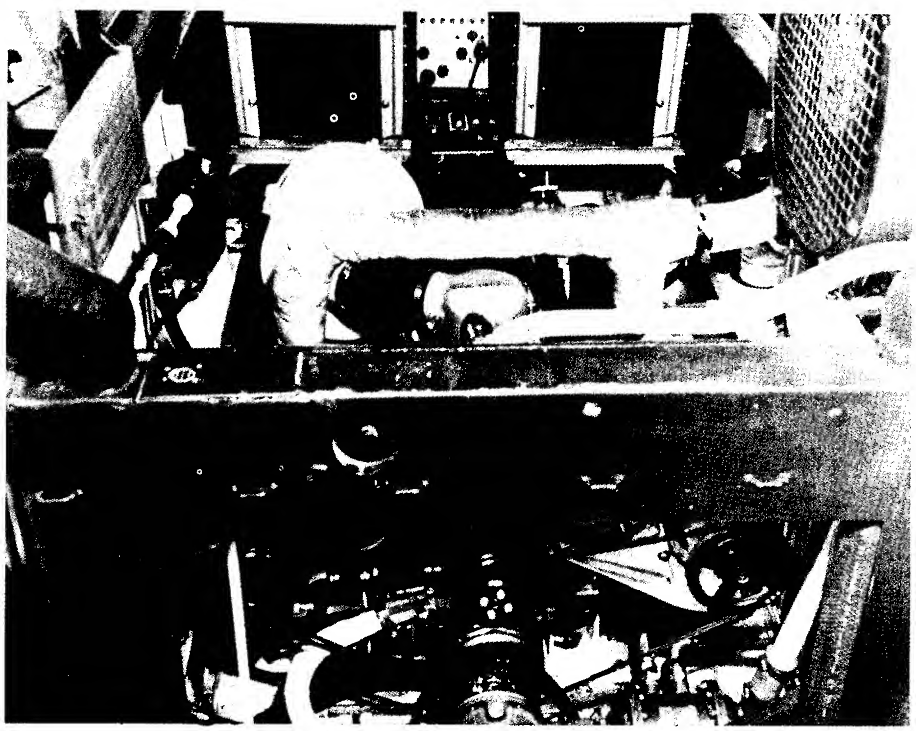


Fig 3. View of Engine Compartment with all access covers open, illustrating accessability for power plant removal

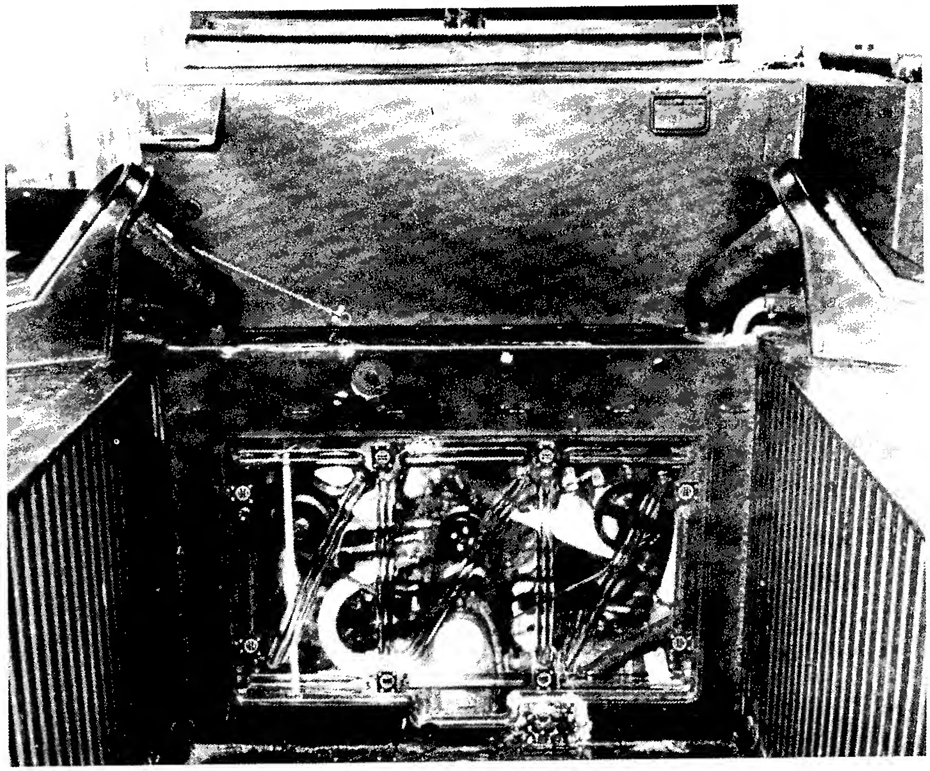


Fig 4. Accessability to Engine Compartment through rear engine bulkhead

SPECIFYING MAINTAINABILITY-DEMONSTRATION-TEST PARAMETERS*

INDEX SERIAL NUMBER — 1090

Harold Balaban

ARINC Research Corporation

SUMMARY

A demonstration procedure is essentially a statistical test of a hypothesis. For maintainability, the hypothesis is generally of the form that a specified maintainability characteristic (e.g., mean active-corrective-maintenance time) meets a specified numerical value. Accordingly, the standard approach has been one of acceptance sampling, in which known risks of wrong decisions (rejecting a satisfactory product or accepting an unsatisfactory product) are considered in relation to sample size (e.g., number of maintenance actions observed).

In this paper we consider the important task of specifying the maintainability characteristics associated with a demonstration test and the corresponding test risks. Specifically, guidelines are provided for determining (a) the type of maintainability index to specify, (b) the acceptable and unacceptable values for this index and, (c) the risks associated with the statistical tests.

REQUISITES FOR A MAINTAINABILITY-DEMONSTRATION-TEST SPECIFICATION

A maintainability-demonstration-test specification is defined here as a set of numerical requirements and associated risk levels that will govern the design and decision criteria of the test. For the most common tests, this specification involves decisions regarding the following:

- Type of maintainability index to be specified
- Acceptable and unacceptable values of the index
- Associated risk levels

For example, the test specification might be as follows:

Null Hypothesis (H_0): Mean corrective-maintenance man-hours = 40 minutes

Alternative Hypothesis (H_1): Mean corrective-maintenance man-hours = 80 minutes

Producer's Risk (α) = 0.20; Consumer's Risk (β) = 0.10

A test based on this specification must be designed such that

$$P\,(\text{reject} \mid \overline{MH}_{ct} = 40 \text{ min}) = 0.20$$

$$P\,(\text{accept} \mid \overline{MH}_{ct} = 80 \text{ min}) = 0.10$$

The following are some of the more important requirements for a maintainability-demonstration-test specification:

- The maintainability index should represent a measure that is directly influenced by equipment design so that the producer can plan for high assurance of a pass decision but bears the responsibility for a reject decision.
- Relationships (at least qualitative) between design parameters and the maintainability index should be known so that design evaluations and predictions are possible.
- The maintainability index should be appropriate for, and measurable in, the demonstration-test environment.
- The maintainability index should be related to higher-level system-requirement parameters, and numerical values should be consistent with values for these higher-level parameters.
- Adequate sampling and statistical evaluation procedures should be available for demonstrating conformance to the requirement.
- Specified maintainability-index and risk values should not lead to sample sizes that exceed available test resources.

Not all of these requisites are necessarily consistent, and often they cannot all be completely satisfied. A requirement consistent with higher-level goals may result in specified values that require sample sizes larger than expected. Tests for conformance to certain types of requirements may require complex statistical procedures, which may not be desirable.

It is, therefore, important that the demonstration-test specification be prepared as early as possible so that its implications can be fully evaluated. This will allow time for a trade-off analysis between test costs and the risks of incorrect decisions.

TYPE OF INDEX

Many different types of indices can be specified for a maintainability demonstration. Some of the standard alternatives for three major factors are as follows:

Factor	Alternative
Type of Maintenance Action	Corrective maintenance, preventive maintenance, total maintenance
Type of Statistical Measure	Mean, median, variance, percentile
Type of Time Measurement	Equipment downtime, man-hours, man-hours per operating hour

On a combinatorial basis, this listing represents a possible 36 alternatives; e.g., one is mean corrective-maintenance man-hours. In addition, there may be multiple parameters such as a mean and percentile, as well as the specification of higher-level indices, such as availability or effectiveness, that embody maintainability.

To provide a guideline for the appropriate choice of an index for demonstration, the selection matrix shown in Figure 1 was developed. To use the matrix, each of the conditions listed at the top of the figure that apply to the equipment of interest should be checked. The appropriate index is then found from the matrix by locating the column that contains an x for each condition checked. For example, if steady-state availability is a critical parameter (Condition 1) and maintenance time is limited by environmental or operational circumstances (Condition 5), the recommended index provides a control on both the mean and maximum maintenance time, and there is an option for including preventive-maintenance time depending on equipment use, scheduling, or criticality.

The set of conditions listed is not exhaustive, but it is believed to include the most important ones.

The major considerations that led to the development of the matrix include the following:

- The mean is directly related to steady-state availability and is therefore the index of choice when this operational requirement exists.
- If the distribution of maintenance times is unknown, the median is preferred since it permits distribution-free tests. If availability is critical, however, use of the central-limit theorem permits a mean test, provided the sample size is large.
- For the lognormal distribution, the median is preferred to the mean (assuming that Condition 2 applies and that 5 and 6 do not) since it is based on only one parameter, which makes statistical analysis exact.
- When maintenance time is limited (Condition 5) the M_{max} index is preferred.
- The mean is preferred over the median if manpower control is also required because the mean is more directly related to man-hours. However, if the distribution is unknown, the median may be used as long as availability is not critical.

*Much of the work on which this paper was based was performed by ARINC Research Corporation for Rome Air Development Center under Air Force Contract F30602-68-C-0047 (Reference 1).

Condition		Condition Identification (Place X in appropriate boxes)
☐	1	Steady-state avilability is a critical parameter.
☐	2	Steady-state availability is not a critical parameter.
☐	3	Maintenance-time distribution is unknown.
☐	4	Maintenance-time distribution is expected to be lognormal.
☐	5	Environmental or operational circumstances limit maintenance time.
☐	6	Manpower allocation or cost is an important factor.

Selection Matrix[1]

Condition \ $\overline{M}$ Index	$\overline{M}_{ct}$ and $\overline{M}_{pt}$[2]				$\tilde{M}_{ct}$			M_{max}			MMH			$\overline{M}_{ct}$ and $M_{max\ ct}$[1] $\overline{M}_{pt}$ and $M_{max\ pt}$				$\tilde{M}_{ct}$ and $M_{max\ ct}$ $\tilde{M}_{pt}$ and $M_{max\ pt}$
1	X	X	X	X										X	X	X	X	
2					X	X	X	X	X	X	X	X	X				X	X
3		X				X			X			X			X			X
4		X	X				X			X			X			X	X	
5								X	X	X				X	X	X	X	X
6				X							X	X	X				X	X

<u>Notation</u> $\overline{M}$ = mean, $\tilde{M}$ = median, M_{max} = maximum maintenance time (percentile),

MMH = maintenance man-hours, ct = corrective maintenance,

pt = preventive maintenance

<u>Notes</u> ① The inclusion of preventive-maintenance indices is optional depending on scheduling and criticality.

② A combined total-maintenance-time index can be used instead of separate indices for corrective and preventive maintenance.

Figure 1. **PROCEDURE FOR MAINTAINABILITY-INDEX SELECTION**

Complete dependence on this procedure is to be avoided. Because of the wide variety of equipments, mission objectives, and environmental and operational circumstances, the selection matrix should be considered a guide only. Ultimately, the best measure is determined by individual system circumstances and good judgement.

SPECIFIED VALUES

The usual specification of values for maintainability demonstration involves assignment of two values for the index selected — a desirable value asociated with the null hypothesis, H_0, and an undesirable (sometimes called marginally acceptable) value associated with the alternative hypotheses, H_1.

In assigning such values, it is reasonable first to consider the goal or H_0 value, since this is what the producer and consumer both seek, and then to assign the H_1 value, which will be a function of the desirable value, minimum operational goals, and required sample sizes.

There are three basic criteria for specifying the desirable values of the selected maintainability index:

(1) The specified value should be consistent with higher-system-level requirements.

(2) It should be realistic.

(3) It should be appropriate to the demonstration environment.

Some simple models for obtaining a maintainability-index requirement that is consistent with a higher-level requirement are reviewed here for two types of availability requirements — point availability and interval availability.

Point Availability

Point availability is the probability that the system is available for operational use at a random point in time. From a long run or steady state viewpoint, it is expressed as the ratio of total "on" time to total time.

When preventive or noncorrective maintenance can be scheduled so that it does not conflict with mission objectives, the following expression is applicable

$$A = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}$$

where

MTBF = mean time between failures

MTTR = mean time to repair

For this case, the simple trade-off relationship $\text{MTTR} = \text{MTBF}\left(\frac{1}{A} - 1\right)$ can be used as the basic model for establishing requirements on MTTR and MTBF.

For systems whose mission is continuous, such as an early warning radar, availability can be expressed by the general steady-state equation

$$A = \frac{\text{MTBM}}{\text{MTBM} + \text{MDT}}$$

where

MTBM = mean time between maintenance

MDT = mean downtime

To develop trade-off relationships, we must consider both preventive and corrective maintenance, but we shall do so at a relatively elementary level. It will be assumed that preventive maintenance is scheduled every T_p hours regardless of when the last maintenance action took place. For example, if the system is a set of light bulbs, preventive maintenance involves complete bulb replacement, and $T_p = 500$, then all bulbs are replaced every 500 hours even if some of them are replaced with relatively little accumulated life.

We also assume that the mean time between failure of the system (θ) is a function of the preventive maintenance period (T_p). In general, this dependency will be denoted by $\text{MTBF} = \theta(T_p)$ where $\theta(T_p)$ will usually be a non-increasing function of T_p. Similarly, we might expect that as T_p increases, the average time required to perform preventive maintenance increases.

To develop a steady-state expression for the mean time between maintenance actions, consider a long time period T. Over T hours, we would expect $T/\theta(T_p)$ corrective maintenance actions and T/T_p preventive maintenance actions. Therefore, the average time between maintenance actions is $T/[T/\theta(T_p) + T/T_p]$ leading to the equation

$$\text{MTBM} = \frac{\theta(T_p)T_p}{\theta(T_p)+T_p}$$

For obtaining mean down-time (MDT), we may expect that an average of $[T/\theta(T_p) + T/T_p]$ actions will have taken place over time period T. Consequently, the probability that a maintenance action is corrective is

$$P[CM] = \frac{T/\theta(T_p)}{T/\theta(T_p)+T/T_p} = \frac{T_p}{\theta(T_p)+T_p}$$

and, similarly,

$$P[PM] = \frac{\theta(T_p)}{\theta(T_p)+T_p}$$

If $\overline{M}_{ct}$ and $\overline{M}_{pt}$ represent the average down-times due to corrective maintenance and preventive maintenance, respectively, we have

$$\text{MDT} = \frac{T_p}{\theta(T_p)+T_p}\overline{M}_{ct} + \frac{\theta(T_p)}{\theta(T_p)+T_p}\overline{M}_{pt}$$

Then for point availability

$$A = \frac{\theta(T_p)T_p}{\theta(T_p)T_p + \theta(T_p)\overline{M}_{pt} + T_p\overline{M}_{ct}}$$

$$= \frac{1}{1 + \overline{M}_{ct}/\theta(T_p) + \overline{M}_{pt}/T_p}$$

Of particular interest for maintainability demonstration is a choice of values for T_p, $\bar{M}_{pt}$, and $\bar{M}ct$ given a requirement on A. If the time interval between preventive maintenance actions (T_p) is increased, it might be reasonable to increase $\bar{M}_{pt}$ since the tasks may be more extensive as a result of the longer operating time. Also θ may be adversely affected if T_p is made too long. On the other hand, too small a value for T_p increases the number of down-times due to preventive maintenance; and while θ may be increased somewhat and $\bar{M}_{pt}$ decreased, there is a minimum T_p value below which it would be unwise to specify.

A general trade-off relationship is difficult to develop because the interrelationships that exist may be varied and complex. Instead, a simple numerical example that illustrates how one may proceed is provided here.

Assume that there is an availability requirement of 0.96. From past experience, feasibility analyses, and operational requirements, the following are reasonable ranges for the parameters listed:

$$\theta = 50 \text{ to } 150 \text{ hours}$$

$$T_p = 25 \text{ to } 75 \text{ hours}$$

$$\bar{M}_{pt} = 1 \text{ to } 3 \text{ hours}$$

$$\bar{M}_{ct} = 1 \text{ to } 4 \text{ hours}$$

If the worst extremes for the means are considered, i.e. $\theta = 50$, $\bar{M}_{pt} = 3$, $\bar{M}_{ct} = 4$, then

$$A = \frac{1}{1 + 4/50 + 3/T_p}$$

$$= \frac{1}{1.08 + 3/T_p}$$

and no positive value of T_p will yield an availability of 0.96. On the other hand, for the best mean values of $\theta = 150$, $M_{pt} = M_{ct} = 1$ we have

$$A = \frac{1}{1 + 1/150 + 1/T_p}$$

and T_p greater than 29 will yield an availability greater than 0.96 indicating that the goal is feasible with an appropriate set of requirements.

Assume now that a more detailed analysis between T_p, θ, and M_{pt} yields the following alternatives:

Alternative	T_p	Max θ	Min $\bar{M}_{pt}$
I	25	150	1.0
II	50	100	1.5
III	75	75	2.0

The value of $\bar{M}_{ct}$ that provides an availability of A is determined from the following equation:

$$\bar{M}_{ct} = \frac{\theta(T_p)}{A}\left[1 - A - \frac{A\bar{M}_{pt}}{T_p}\right]$$

The results for A = 0.96 are as follows:

Alternative	Mct
I	0.25
II	1.17
III	1.13

Because there is an initial restriction on $\bar{M}_{ct}$ of $1 \leq \bar{M}_{ct} \leq 4$, Alternative I cannot be chosen. Therefore, the choice is between II and III, and this decision depends on the costs associated with the specific values of T_p, θ, $\bar{M}_{pt}$, and $\bar{M}_{ct}$.

This particular example involves the selection of a preventive-maintenance schedule as well as mean corrective-maintenance and preventive-maintenance times. Much more sophisticated models for preventive-maintenance scheduling have been developed, and in practice the procedure might be to use one of these models to select T_p and θ and then choose values for $\bar{M}_{ct}$ and $\bar{M}_{pt}$ to meet the availability goal.

Interval Availability

Interval availability is the probability that the system will be available for operational use within a specified time interval. It is applicable when the system is required to perform a series of missions; the most common example of such a system is an aircraft. For such cases, it is often important to use an interval-availability requirement to control the probability of readiness after completion of a mission.

A model for this type of requirement can be fairly complex depending on the system, operational conditions, and assumptions made. A relatively simple model for steady-state interval availability is presented below. It assumes a Markov process for the mission/service-repair sequence, a constant mission time T, and a constant allowable repair time t.

The following four functions will be considered:

A(t) — the probability that the system is available within t hours after the scheduled mission is completed

R(T) — the system reliability for a mission of T hours

S(t) — the probability that necessary servicing (e.g., refueling and rearming an aircraft) is performed within t hours after a successful mission

M(t) — the probability that servicing and any necessary repairs can be accomplished within t hours after maintenance is initiated on a failed system

The steady-state interval availability is then given by the following equation [a bar above a symbol represents the complementary event, e.g., $\bar{R}(t) = 1 - R(t)$]:

$$A(t) = A(t)R(T)S(t) + A(t)\bar{R}(T)M(t) + \bar{A}(t)M(T+t)$$

The first term on the right-hand side is the probability that the system was available at the start of the previous mission, did not fail in T hours of operation, and is serviced within t hours. The second term represents the probability that the system was available at the start of the previous mission, that a failure occurred during that mission, and that repair and servicing are completed within t hours. The third term is the probability that the system was unavailable at the start of the previous mission and that repair and servicing are completed before the start of the current mission (a total time of T+t hours).

Solving for A(t) yields

$$A(t) = \frac{M(T+t)}{1 - R(T)S(t) - \bar{R}(T)M(t) + M(T+t)}$$

The maintainability parameters of interest are M(t), M(T+t), and S(t). M(T+t) should equal 1 since this represents the probability that maintenance is completed within the usual allowable time (t) plus the mission time T. Then

$$A(t) = \frac{1}{2 - R(T)S(t) - \bar{R}(T)M(t)}$$

Since a maximum of t hours is available for servicing and corrective maintenance, servicing should be completed in much less time than t hours to permit corrective maintenance to take place. In this case, a time $t_s < t$ can be chosen such that requirements are to be placed on $S(t_s)$ and $M(t_c)$, where t_s plus t_c is less than or equal to t. $S(t_s)$ equals the probability that servicing is completed within t_s hours, and $M_c(t_c)$ equals the probablity that corrective maintenance is completed within

time t. Then M(t) can be replaced by $S(t_s) \times M_c(t_c)$ (assuming the independence of the two associated events). The use of this product is conservative since it is assumed that only tc hours are available for corrective maintenance even if servicing is completed earlier than ts hours. The availability model is then

$$A(t) = \frac{1}{2 - R(T)S(t_s) - \bar{R}(T)S(t_s)M_c(t_c)}$$

Again, cost and operational factors will determine which of the appropriate combinations of R, $S(t_s)$, and $Mc(t_c)$ should be specified for a given availability requirement. In this example the $S(t_s)$ and $M_c(t_c)$ requirements are often called M_{max} type requirements, which are actually percentile values of the cumulative distribution function. Figure 2 illustrates the trade-off relationship when R(T) is fixed at 0.95.

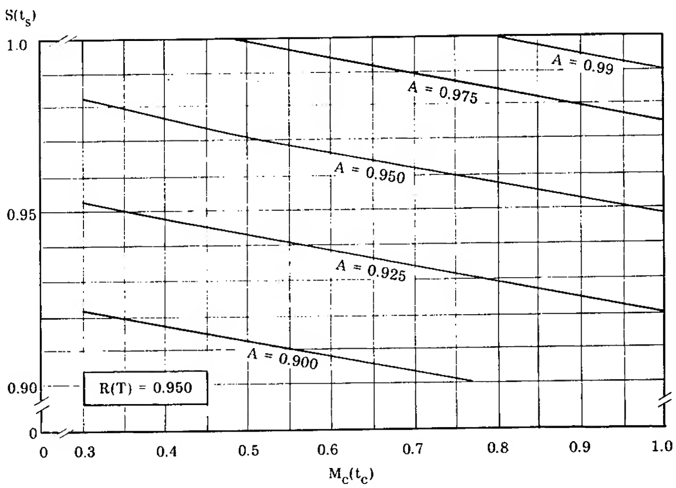


Figure 2. **TRADE-OFF RELATIONSHIPS FOR INTERVAL AVAILABILITY**

The above-described approaches for obtaining maintainability requirements from an overall availability requirement are only indicative of the type of approach that can be used. Several simplifying assumptions were made in establishing the relationships, some possibly important factors were not included, and cost was considered only qualitatively. Therefore, the equations and curves presented for obtaining specified values must be adjusted to account for factors that have not been considered adequately in this general model.

REALISM OF SPECIFIED VALUES

Approaches similar to those presented above make it possible to specify a maintainability value. The next consideration is realism. It is necessary first to establish what is meant by a realistic value. Expressions such as "within the state of the art" are commonly encountered, and while they do not provide a quantitative assessment, they do convey the general belief that the value can be achieved by current technological capability.

Since maintainability-demonstration-test requirements must be established very early in the development program (often before contract award), the most logical approach to assessing realism — and sometimes even establishing the requirement if allocation from higher levels is not required — is to evaluate the maintainability performance of existing systems similar to that under development. If the basic maintainability design is known at the time the requirement is to be established, an applicable prediction technique can be exercised.

Whether historical data or prediction, or both, are used for assessing realism, careful judgment is required. If an allocation leads to an M_{ct} value of 20 minutes but a 30-minute value was observed for the most similar existing system, can it be concluded that 20 minutes is unrealistic? The following questions must be considered:

- How similar are the items?
- How similar will the maintenance environment be?
- Since the observed 30-minute value is necessarily based on a sample, what is the lower confidence limit associated with such a mean-value estimate?
- How much maintainability improvement can reasonably be asked for?
- Is there any margin for increasing the 20-minute specified value?

Again, the answers to these questions and the conclusions to be drawn depend on individual circumstances. To check for realism, prediction techniques such as those presented in References 2 and 3 can be used as applicable.

Observed maintainability values of existing equipments obtained from several sources are presented in Reference 1 to provide historical data that can be used as a guide in assessing the realism of a specified value.

APPLICABILITY OF REQUIREMENTS TO THE DEMONSTRATION ENVIRONMENT

In a maintainability-demonstration survey (Reference 1) it was found that a frequently cited difficulty was the difference between test environment and field environment. In an RADC study*, a comparison of demonstration-test results with field operational results for seven systems revealed wide discrepancies. The operational field MTTR was always greater. Although the field data may have been contaminated with undesirable factors, such as administrative-time delays, the observed differences are still quite illuminating.

It is apparent that the closer the test environment is to the expected field environment, the more meaningful the demonstration test will be, and that every effort should be made to achieve such similarity. Specific reasons for biases due to test environment are outlined in this section.

Unless an operational-type test is to be performed, demonstration environments will differ in some respects from the field environment. Because such differences do exist, a maintainability-demonstration requirement based on operational goals should not be applied unless its applicability to the demonstration conditions is first considered.

As a general principle, the specified value based on operational goals and conditions must be suitably adjusted to reflect the maintenance environment governing the demonstration. Often, this is a difficult principle to adhere to. With an avionic equipment, for example, a certain amount of time will be spent in the field just reaching the equipment in the aircraft, and the time to locate the malfunction and complete repairs and checkout is a function of this accessibility factor. If the demonstration test is not to take place in the aircraft (and this is often the case), there is the question of whether the specified value should be adjusted, and by how much.

It might be possible to construct a mockup to simulate the actual conditions, thus eliminating the need for adjustment. Generally, this type of simulation will not be possible, and field and test conditions must be carefully analyzed and their effects quantitatively assessed. Table 1 lists various factors to be considered in evaluating the applicability of a specified maintainability index. Table 2 lists some specific causes of discrepancies that are classified as yielding either pessimistic or optimistic results.

RISK ASSIGNMENT

There are generally two risks involved in a demonstration test:

(1) Producer's risk, α — the probability of rejection if the maintainability characteristic is at the desired level.

(2) Consumer's risk, β — the probability of acceptance if the maintainability characteristic is at the minimum acceptable (or undesirable) level.

Ideally, α and β would be equal to zero. Granting that this is impossible, very small values of α and β — on the order of 0.001 — are desirable. Such small values are impractical, however, since the selection of α and β associated with the H_0 and H_1 values for maintainability dictates the sample size. For α and β on the order of 0.001, sample sizes far exceeding available test resources will usually be required.

*A. Coppola and J. Deveau, "Reliability and Maintainability Case Histories", *Annals of Reliability and Maintainability*, Vol. 6, 1967, pp. 582—586.

Table 1. FACTORS AFFECTING THE SUITABILITY OF A SPECIFIED MAINTAINABILITY INDEX FOR MAINTAINABILITY DEMONSTRATION	
Physical Equipment	Training and experience
Stage of completion	Indoctrination
Similarity to production items	**Support Items**
Physical location	Tools
Interfacing equipment	General and special test equipment
Test Location and Facility	Spares availability
Lighting factors	Technical manuals
Weather factors	**Operational Factors**
Space factors	Mode of equipment operation
Test Team	Procedures for instituting maintenance
Organization	Procedures for fault selection

Table 2. CAUSES OF DISCREPANCIES BETWEEN TEST RESULTS AND FIELD RESULTS

Causes of Optimistic Test Results

- The demonstration maintenance technicians are not representative of typical maintenance personnel because they have more education and training or greater knowledge of the equipment design.
- The monitoring situation imparts to the technician an urgency not normally encountered in the field.
- Known probable tasks are rehearsed beforehand.
- Necessary support equipment is readily available.
- Observed times are not contaminated with such factors as administrative or logistic delay, as field results sometimes are.
- Difficult-to-isolate faults such as intermittencies and degradation failures are not simulated.

Causes of Pessimistic Test Results

- The technicians are not familiar with the equipment and have not acquired the necessary experience for rapid fault isolation.
- Field and procedural modifications to reduce maintenance time have not yet been made.
- Initial manuals may be incomplete or require revision.
- The monitoring situation can adversely affect the technician's performance.

For example, consider a test of the mean of a lognormal distribution such as the following:

$$H_0: \quad \mu = \mu_0 = 30 \text{ minutes}$$

$$H_1: \quad \mu = \mu_1 = 45 \text{ minutes}$$

As shown in Reference 1, the necessary sample size for this test is given by the equation

$$n = \frac{\left(Z_\alpha \mu_0 + Z_\beta \mu_1\right)^2}{(\mu_1 - \mu_0)^2} \left(e^{\sigma^2} - 1\right)$$

where Z_α and Z_β are the normal deviates corresponding to the $(1 - \alpha)$th and $(1 - \beta)$th percentile of a normal (0, 1) distribution and σ^2 is the variance of the logarithm of maintenance time. If $\alpha = \beta$ and $\sigma^2 = 1$ are assumed, then

$$n = \frac{Z_\alpha^2 \ (30 + 45)^2}{(45 - 30)^2} (e^1 - 1) = 43 Z_\alpha^2$$

From this equation it can be shown that if $\alpha = \beta = 0.10$, 70 observations are required. If α and β are reduced to 0.01, about 230 observations are necessary; and for $\alpha = \beta = 0.001$, a sample size of more than 400 is called for.

Most development budgets and schedules will not allow for a test requiring 400 sample observations even if the failures are to be simulated. In fact, even a sample size of 70 may tax available rsources, and for this illustrative case, risks on the order of 0.15 or 0.20 may be necessary.

It is not necessary, of course, for α to equal β. If, for example, the need for the equipment is great and a 45-minute mean time to repair can be tolerated (perhaps with later improvement by modification and appropriate training, manning, and support planning), there is a relatively low risk of rejecting good equipment and a higher risk of accepting a minimally acceptable equipment.

Use of Prior Information in Risk Trade-Off

The choice of α and β is also one involving trade-offs. From a decision-theory viewpoint, the trade-off can be normalized to a cost criterion based on the following factors:

(1) Cost of testing (sample size)

(2) Cost of rejecting good equipment

(3) Cost of accepting poor equipment

While the first factor can generally be costed in terms of manpower, facilities, and time, the second and third factors are more difficult to assess quantitatively. Assuming that prior information is available for estimating at least relative values associated with the three costs, two simplified approaches employing decision-theory concepts for selecting α and β are discussed below. For convenience, the maintainability characteristic of interest will be denoted by M, and specified H_0 and H_1 values by M_0 and M_1, respectively. Also let

$$C_0 = \text{Cost of rejection if } M = M_0$$

$$C_1 = \text{Cost of acceptance if } M = M_1$$

Minimax Criterion

The minimax criterion is used when it is desirable to avoid extremely high costs. To use this criterion, for a given combination of α and β, say (α_i, β_j), compute the following:*

$$(1) \quad L_{ij}(M_0) = C_0 \alpha_i + C_{ij}(M_0)$$

$$(2) \quad L_{ij}(M_1) = C_1 \beta_j + C_{ij}(M_1)$$

$$(3) \quad L_{ij} = \text{Max}\left[L_{ij}(M_0), L_{ij}(M_1)\right]$$

where

$C_{ij}(M_k)$ = Test costs associated with (α_i, β_j) if $M = M_k$ (k = 0 or 1)

$L_{ij}(M_k)$ = Total cost if $M = M_k$ (k = 0, 1) and $\alpha = \alpha_i$, $\beta = \beta_j$

L_{ij} = Maximum cost if $\alpha = \alpha_i$, $\beta = \beta_j$

Generally $C_{ij}(M_k)$ will be a function of the sample-size requirements dictated by the α_i, β_j pair and will not depend on M except for sequential tests, for which the average value of n given $M = M_k$ can be used.

The α, β risk pair to select is that which has the minimum value of L_{ij}. By this criterion the selected risks are such that the maximum possible costs are minimized.

As an example of this procedure, consider the illustrative test discussed above. For simplicity, assume that the values of α and β to be considered are restricted to 0.05, 0.10, 0.20. The possible risk pairs and associated sample sizes, from the previous equation, are as follows:

Pair (i, j)	α	β	n_{ij}
11	0.05	0.05	116
12	0.05	0.10	87
13	0.05	0.20	58
21	0.10	0.05	96
22	0.10	0.10	70
23	0.10	0.20	44
31	0.20	0.05	75
32	0.20	0.10	52
33	0.20	0.20	30

*These equations are based on the assumption that no costs except test costs are associated with an accept decision if $M=M_0$, or with a reject decision if $M=M_1$.

Costs considerations lead to the following relationships:

$$C_0 = \$50{,}000$$

$$C_1 = \$40{,}000$$

$$C_{ij} = \$2{,}000 + n_{ij}^2$$

The results of the necesary computations are shown in Table 3. For each pair, the maximum value of L_{ij} is underscored. The minimum of these maximum values is seen to be $11,900, which is yielded by the pair $\alpha = 0.10$, $\beta = 0.10$.

Table 3. COMPUTATIONS FOR OBTAINING MINIMAX RISKS FOR ILLUSTRATIVE EXAMPLE

Index		Risks		Costs	
i	j	α	β	$L_{ij}(M_0)$	$L_{ij}(M_1)$
1	1	0.05	0.05	$17,956	$17,456
1	2	0.05	0.10	12,069	13,569
1	3	0.05	0.20	7,864	13,364
2	1	0.10	0.05	16,216	13,216
2	2	0.10	0.10	11,900*	10,900
2	3	0.10	0.20	8,936	11,936
3	1	0.20	0.05	17,625	9,625
3	2	0.20	0.10	14,704	8,704
3	3	0.20	0.10	12,900	10,900

*Minimum of maximum values.

Bayes Strategy

For the Bayes approach, prior information or subjective evaluation is required to estimate the following:

$$P_0 = \text{probability } M = M_0$$

$$P_1 = 1 - P_0 = \text{probability } M = M_1$$

Then for each pair (i,j) the expected cost is computed:

$$E_{ij} = P_0 \left[C_0\alpha_i + C_{ij}(M_0)\right] + P_1 \left[C_1\beta_j + C_{ij}(M_1)\right]$$

The pair for which E_{ij} is a minimum is selected.

In this procedure, the risks are selected to minimize the expected costs.

To illustrate this procedure, assume that it can be reasonably estimated from past performance data, in conjunction with evaluation of the maintainability-program efforts, that $P_0 = 0.70$, $P_1 = 0.30$. The values associated with this prior distribution are shown in Table 4.

Table 4. EXPECTED COSTS FOR ALTERNATIVE PLANS

Index		Risks		Expected Costs,
i	j	α	β	E_{ij}
1	1	0.05	0.05	$ 17,806
1	2	0.05	0.10	12,519
1	3	0.05	0.20	9,514*
2	1	0.10	0.05	15,316
2	2	0.10	0.10	11,600
2	3	0.10	0.20	9,836
3	1	0.20	0.05	15,225
3	2	0.20	0.10	12,904
3	3	0.20	0.20	12,300

*Minimum value.

From this listing, it is seen that the risk $\alpha = 0.05$, $\beta = 0.20$ minimizes expected cost. If the prior probabilities were $P_0 = 0.50$, the pair $\alpha = 0.10$, $\beta = 0.20$ would be optimal. With the prior estimates of P_0 and P_1, the expected cost without testing can also be evaluated. If no testing is performed and the equipment is to be accepted upon delivery. the expected cost is simply

$$(P_1)(C_1) = (0.30)(40{,}000) = \$12{,}000$$

For this example, the decision not to test is unwise. However, where testing is quite costly and past performance indicates a high probability of a satisfactory product, this type of evaluation might indicate that, from the viewpoint of economy, little or no testing is the preferred choice.

Discussion of Decision-Theory Approaches

The two decision-theory approaches described above. might be criticized on the basis that only the H_0 and H_1 values for M are considered. More extensive procedures can be used, but they require prior information and cost relatonships that are not generally available.

In defense of the procedure, it can be said that for conventional sampling procedures, in which α and β are more or less arbitrarily chosen, two levels of maintainability are also considered. Moreover, the M_0 and M_1 values and their associated risks do determine the complete operating-characteristic curve. Choosing α and β from a decision-theory viewpoint does provide some cost control for the test procedure and thus has an economic advantage over nondecision-theory approaches.

CONCLUSION

The procedures outlined in this paper for specifying a maintainability-demonstration-test requirement consider the important areas of index selection and appropriate levels of specified maintainability and test risks. Criteria relating to applicability, realism, and economics were applied in developing the guidelines and models. Although a particular equipment/mission application may require more complex procedures than those presented in this paper, the same general critera should apply. We must also note that the benefits derived from a carefully developed procedure for specifying maintainability demonstration test parameters can be quickly lost if equal consideration is not given to the management planning, sampling, statistical analyses, and administration aspects of the demonstration test. References 1, 3, 4 and 5 consider these aspects in some detail.

REFERENCES

1. H. Balaban, *Maintainability Prediction and Demonstration Techniques, Volume II — Maintainability Demonstration*, ARINC Research Corporation, A.F. Contract F30602-68-C-0047, RADC TR-69-356, Vol. II, January 1970.

2. G. Griswold, et al., *Maintainability Prediction and Demostration Techniques* [*Final Technical Report*], ARINC Research Corporation, A.F. Contract F30602-68-C-0087, RADC TR-70-89, June 1970.

3. MIL-HDBK 472, *Maintainability Prediction.*

4. H. Balaban, "Statistical Test Plans for Maintainability Demonstration", *Proceedings of the 1972 Annual Reliability and Maintainability Symposium*, January 1972, pp. 386—382.

5. MIL-STD 471, *Maintainability Demonstration.*

PROBABILISTIC ANALYSIS OF SEQUENTIAL TEST PLANS

B. Tiger, M. Raphelson, G. Ashendorf
RCA Corporation
Camden, New Jersey

SUMMARY

This paper reports the details of an analysis procedure to determine the decision probabilities which develop during sequential test plans. Currently only the final decision probabilities for an overall sequential test plan are generally described in the literature. How they exactly develop during the testing is not commonly understood. Computer simulations have been reported which estimate, for some specific sequential test plans, the probabilities of the decision alternatives during the testing, but this paper contributes an exact analysis approach applicable to any type of sequential test plan.

The significance of such probabilistic analysis is that a much clearer comprehension is achieved of how an accept decision, reject decision or continue test decisions occur along with their respective probabilities of occurrence as a function of test time.

Computer printouts showing the results for Sequential Test Plans III and IV defined in MIL-STD-781B are presented herein.

INTRODUCTION

Sequential test plans are commonly used to demonstrate equipment and system reliability requirements. MIL-STD-781B contains a selection of sequential test plans. Many systems development programs require successful completion of a selected plan from this standard. It is, therefore, vital for management to thoroughly understand the risks associated with these test plans in order that appropriate program and cost decisions be made. This paper contributes to such greater understanding because the decision alternatives as they develop during the testing, along with their respective occurrence probabilities, are detailed in a more easily understood manner than currently described in the literature.[2,3,4,5]

OBJECTIVE

The purpose of this study is to perform probabilistic analysis of sequential test plans such as defined in MIL-STD-781B. Each of these plans contains explicit criteria for making:

1. An accept decision or

2. A reject decision or

3. A decision to continue testing

as a function of total number of failures and total time under test.

Probabilistic analysis consists of determining the:

1. Probability of an accept decision

2. Probability of a reject decision

3. Probability of a decision (i.e., the sum of the above two probabilities)

as a function of total time under test.

ANALYSIS

The nature of the MIL-STD-781B sequential test plans is illustrated in the first 5 columns of Table 1, in which Test Plan III is described. Column 2 shows time (i.e. T) in multiples of the Mean Time Between Failures specified in a contract or equipment specification (i.e. θ_0 as described in MIL-STD-781B). If 3 failures occur on or before .35 specified MTBF multiples, testing is terminated with a "Reject" decision. The "Reject" numbers are shown in column 3. If 0, 1 or 2 failures have occurred at T = .35, testing is continued. The "Continue Test" ranges are shown in column 4 and the "Accept" numbers (i.e., A) are shown in column 5. If a fourth failure occurs on or before T = 1.04, testing is terminated with a "Reject" decision. If 0 to 3 failures have occurred by T = 1.04, testing is continued. When T reaches 2.20, an "Accept" decision becomes possible; Accept if 0 failures, Reject if a 6th failure has occurred and continue testing if there are 1, 2, 3, 4 or 5 failures. This process can continue until T = 10.30, at which time an "Accept" decision is made if there are 15 or less failures and a "Reject" decision is made as soon as a 16th failure occurs.

Test Plan III is more fully visualized via a System State Phase Model such as described in Reference 6. There are 26 phases. Each phase corresponds to a row in Table 1. The first phase is from T = 0 to T = .35. The second phase is from T = .35 to T = 1.04 and so on until the 26th phase which covers from T = 9.83 to T = 10.30. The first, second and 26th phases are illustrated herein as Figures 1, 2 and 3. Figure 1 identifies all possibilities in the first phase:

1. no failures

2. 1 failure

3. 2 failures

4. other (i.e., 3 or more failures)

The first 3 events imply "continue testing"; the 4th event results in immediate rejection (i.e., end of testing with a decision of inadequate reliability) as soon as the 3rd failure occurs.

The exit states of one phase are the entry states of the next phase. Figure 2 covers the phase from T = .35 to T = 1.04. At T = .35, none, 1 or 2 failures have occurred. Between T = .35 and T = 1.04, testing is terminated with a reject decision as soon as a 4th failure occurs. Therefore, the number of failures which must occur during the phase for a reject decision depends on the number of failures which have occurred by the beginning of the phase. If none occurred, then 0, 1, 2, 3 and 4 or more must be evaluated in Figure 2. If 1 occurred, then 0, 1, 2 and 3 or more must be evaluated. If 2 occurred, then 0, 1 and 2 or more must be evaluated. The number of possibilities at the end of phase 2 are:

1. no failures

2. 1 failure

3. 2 failures

4. 3 failures

5. 4 or more failures

The first four possibilities imply "continue testing" and are the input events to the next phase. The 5th possibility results in the termination of testing with a decision to reject. This process continues for a total of 26 phases.

Such System Phase Models[6] identify all events, their combinations and the consequences of their respective occurrences. Therefore, the model described herein is completely defined by the T, R and A values shown in Table 1. Any appropriate probability distribution can be used, in each phase, to determine the occurrence probability of each respective number of failures. Thus, while MIL-STD-781B presumes a Poisson process and the Wald type Sequential Probability Ratio Test,[7] the SSPM can handle any type of sampling plan and any underlying probability distribution(s).

The probability of occurrence of each possible outcome of a phase (e.g., none, 1, 2, 3 or Reject at the end of Phase 2) is the sum of products of probabilities of each input event times number of failures during the phase which result in each respective outcome. This calculation procedure is illustrated via the enclosed calculations for the first 2 phases of the Test Plan III SSPM model. Here a Poisson distribution and MTBF $=\vartheta_0$ are assumed.

The first phase shows a .006 probability of rejection and the following exit state probabilities:

Probability (no failures)	= .705
Probability (1 failure)	= .246
Probability (2 failures)	= .043
Total	= .994

The probability of rejection is:

Prob Rejection = 1 - .994 = .006

Each exit state probability is appropriately used in Figure 2 as an entrance probability. This Figure shows a .019 rejection probability and the other exit probabilities become entrance probabilities to the following phase.

Results are compiled as shown in the last 3 columns of Table 1. The numerics shown correspond to the calculations of the first 4 phases (i.e., Poisson distribution and MTBF $=\vartheta_0$ are assumed). The rejection probability, for the first row, is also the decision probability (i.e., .006). The reject probability shown in Figure 2 (i.e., .019) adds to the .006 to yield a .025 probability of rejection by T = 1.04. Phase 3 calculates to a .018 reject probability; thus, there is a 0.43 shown in the third row. Phase 4 has a .110 Accept probability. This numeric shows up in the 4th row under Probability of Acceptance. Acceptance was not permitted until T = 2.20. The reject probability is .007 in Phase 4; this adds to the .043 to yield a .050 Probability of Rejection by T = 2.20. The probability of a decision is the sum of the respective accept and reject probabilities; it is the sum of .110 and .050 (i.e., .160) for the 4th row. Remaining calculations are shown via a computer printout in Table 2.

While this calculation procedure is straightforward, it is long and tedious. Proper analysis requires that the entire calculation procedure be performed over a range of assumed values of MTBF; the last 3 columns of Table 1 must be recalculated for each assumed true MTBF value. Therefore, a computer program was prepared to expedite such calculations.

RESULTS

A computer printout of the results for Test Plans III and IV of MIL-STD-781B is shown in Tables 2 and 3. Space limitation in these proceedings prevents presentation of more of the printouts.

CONCLUSION AND RECOMMENDATIONS

The probabilistic analysis presented herein consists of determining the respective decision probabilities throughout the testing. Such analysis is new and has several important ramifications for developing further analytical capability. These include:

1. More rigorous and sensitive analytical models for describing the reliability and effectiveness of complex systems which must perform over a detailed mission profile consisting of phases of different environmental conditions and usage requirements.

2. New statistical tools for estimating universe parameters from sequential test results. Current techniques do not permit rigorous estimation because sequential plans are described only as Tests of Hypothesis.

REFERENCES

1. MIL-STD-781B, "Reliability Tests: Exponential Distribution," 15 November 1967

2. Aroian, L. A., "Exact Truncated Sequential Test for the Exponential Density Function" Ninth National Symposium on Reliability and Quality Control, 1963

3. Epstein, B., Patterson, A. A., Qualls, C. R., "The Exact Analysis of Sequential Life Test With Particular Application to AGREE Plans," Aerospace Reliability and Maintainability Conference, 1963

4. Sumerlin, W. T., "Confidence Calculations for MIL-STD-781," 1972 Annual Reliability and Maintainability Symposium

5. Kremp, B. F., Lupo, J. E., "MIL-STD-781 Sequential Tests: Time to Completion," 1971 Annual Symposium on Reliability

6. Tiger, B., "Evaluating System States in Their Mission Phases," 1969 Annual Symposium on Reliability

7. Wald, A., "Sequential Analysis," John Wiley and Sons, New York (1947)

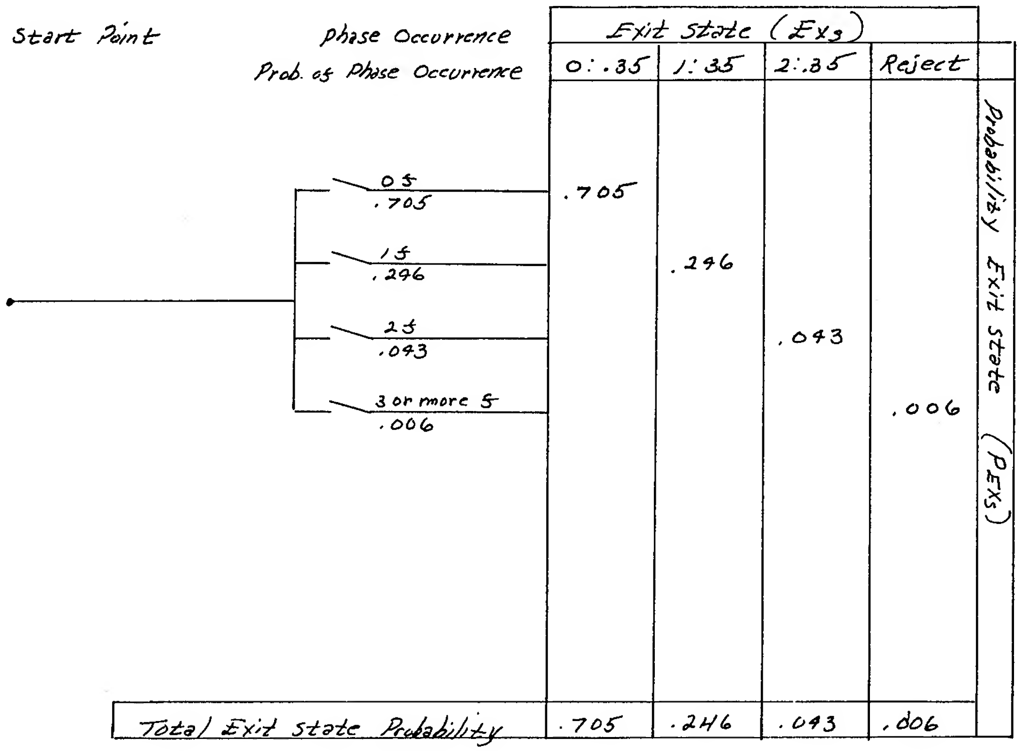
ACKNOWLEDGEMENT

Mr. C. Riggi, of RCA Camden, New Jersey, prepared the computer program which was needed to perform the calculations described herein.

TABLE 1. PROBABILISTIC ANALYSIS: MIL-STD-781B TEST PLAN III

	DEFINITION OF TEST PLAN				PROBABILITY ANALYSIS TRUE MTBF $=\theta_o$		
PHASE	MULTIPLES OF SPECIFIED MTBF (θ_o) ≡ T	REJECT IF NUMBER OF FAILURES BELOW OCCUR ON OR BEFORE TIME IN COLUMN 2 ≡ R.	CONTINUE TEST IF NUMBER OF FAILURES FAIL IN RANGE BELOW AT TIME IN COLUMN 2	ACCEPT IF NO MORE THAN NUMBER OF FAILURES BELOW OCCUR BY TIME IN COLUMN 2 ≡ A	PROBABILITY OF ACCEPTANCE BY INDICATED TIME	PROBABILITY OF REJECTION BY INDICATED TIME	PROBABILITY OF DECISION BY INDICATED TIME
1	.35	3	0 - 2	Not Applicable	0	.006	.006
2	1.04	4	0 - 3	Not Applicable	0	.025	.025
3	1.74	5	0 - 4	Not Applicable	0	.043	.043
4	2.20	6	1 - 5	0	.110	.050	.160
5	2.43	6	1 - 5	0			
6	2.89	7	2 - 6	1			
7	3.12	7	2 - 6	1			
8	3.59	8	3 - 7	2			
9	3.82	8	3 - 7	2			
10	4.28	9	4 - 8	3			
11	4.51	9	4 - 8	3			
12	4.97	10	5 - 9	4			
13	5.20	10	5 - 9	4			
14	5.67	11	6 - 10	5			
15	5.90	11	6 - 10	5			
16	6.36	12	7 - 11	6			
17	6.59	12	7 - 11	6			
18	7.05	13	8 - 12	7			
19	7.28	13	8 - 12	7			
20	7.75	14	9 - 13	8			
21	7.97	14	9 - 13	8			
22	8.44	15	10 - 14	9			
23	8.67	15	10 - 14	9			
24	9.13	16	11 - 15	10			
25	9.83	16	12 - 15	11			
26	10.30	16	-	15			

Calculations by Computer Printout



Sample Calculation

$\lambda = \frac{1}{\theta_0}$

$\Delta T = .35\,\theta_0$

$$PROB\{K \text{ FAILURES}\} = \frac{e^{-\lambda \Delta T} (\lambda \Delta T)^K}{K!}$$

$$PROB\{0\} = e^{-.35} = .705$$

FIG 1
Phase One
From $T=0$ to $T= .35\,\theta_0$

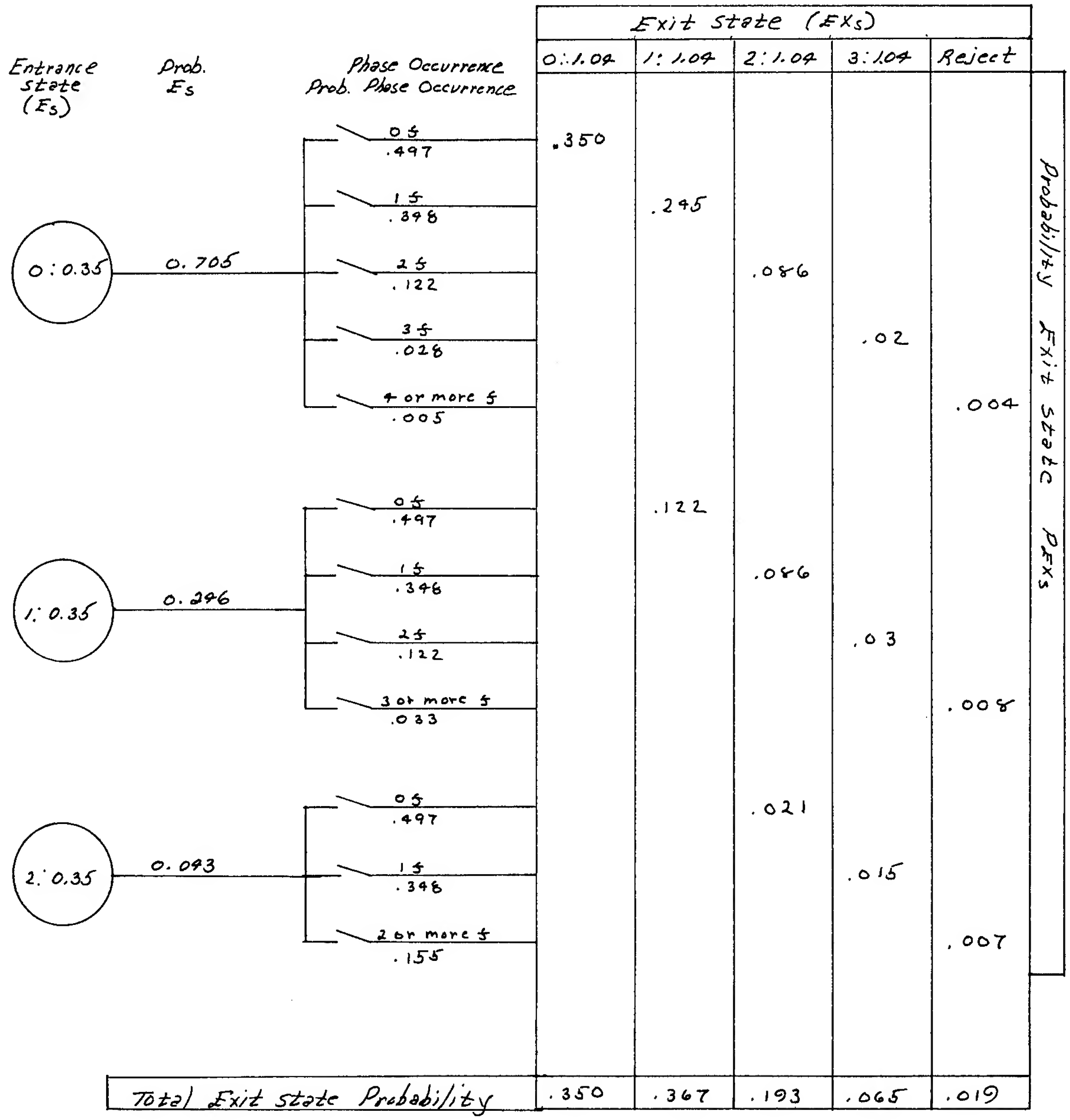


FIG. 2
Phase Two
From $T = .35\,\theta_0$ to $T = 1.04\,\theta_0$

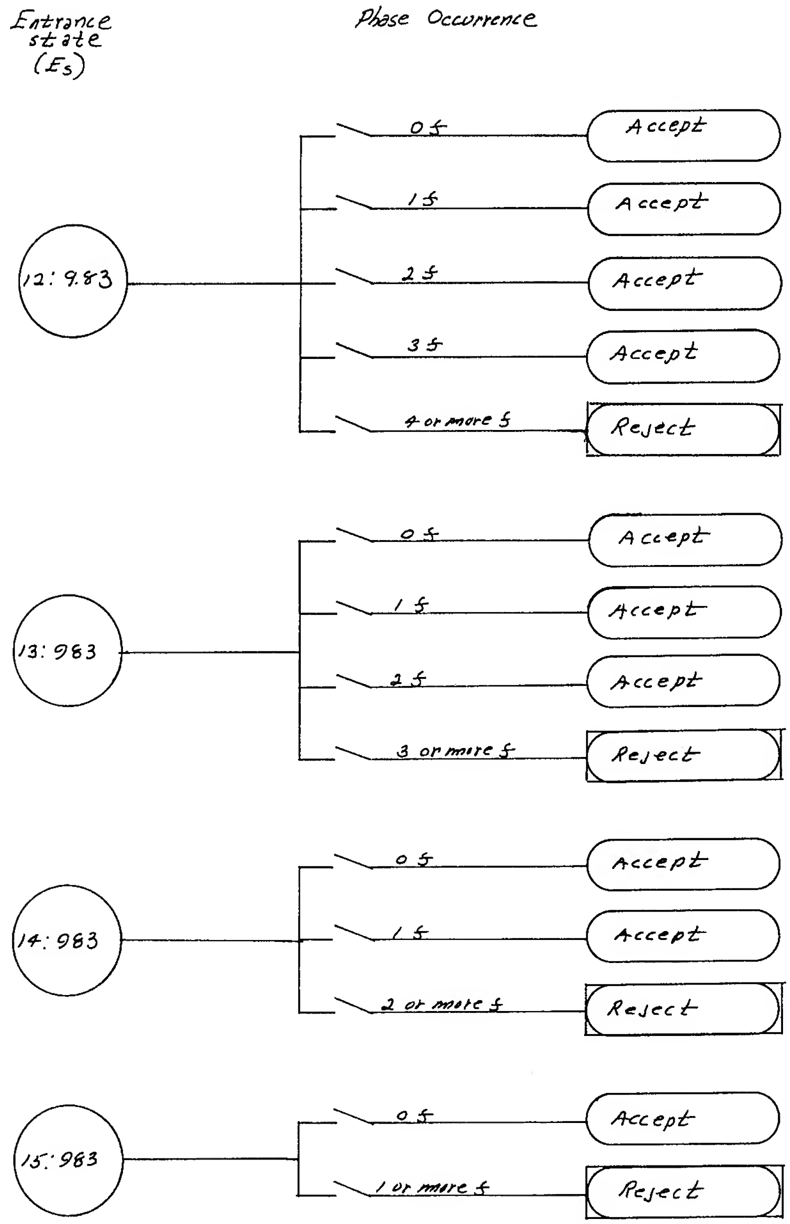


Fig 3
Phase 26
From $T = 9.830\,\theta_0$ to $T = 10.30\,\theta_0$

TEST PLAN III

Test Parameters				True MTBF = $.5\theta_o$			True MTBF = $.75\theta_o$			True MTBF = θ_o			True MTBF = $2\theta_o$		
Time in Multiples of θ_o T	Number of Failures causing Reject	Number of Failures to Continue Test	Number of Failures Allowing Accept	Prob. of Acceptance	Prob. of Rejection	Prob. of Decision	Prob. of Acceptance	Prob. of Rejection	Prob. of Decision	Prob. of Acceptance	Prob. of Rejection	Prob. of Decision	Prob. of Acceptance	Prob. of Rejection	Prob. of Decision
0.350	3	0-2	Not App.	0.000	0.034	0.034	0.000	0.011	0.011	0.000	0.005	0.005	0.000	0.000	0.000
1.040	4	0-3	Not App.	0.000	0.164	0.164	0.000	0.056	0.056	0.000	0.024	0.024	0.000	0.002	0.002
1.740	5	0-4	Not App.	0.000	0.299	0.299	0.000	0.103	0.103	0.000	0.042	0.042	0.000	0.003	0.003
2.200	6	1-5	0	0.012	0.354	0.366	0.053	0.122	0.176	0.110	0.049	0.160	0.332	0.004	0.336
2430	6	1-5	0	0.012	0.409	0.422	0.053	0.145	0.198	0.110	0.058	0.169	0.332	0.004	0.337
2.890	7	2-6	1	0.025	0.454	0.480	0.116	0.161	0.277	0.233	0.063	0.296	0.592	0.004	0.596
3.120	7	2-6	1	0.025	0.499	0.525	0.116	0.180	0.297	0.233	0.070	0.304	0.592	0.004	0.597
3.590	8	3-7	2	0.037	0.537	0.575	0.175	0.195	0.370	0.341	0.075	0.417	0.755	0.005	0.760
3.820	8	3-7	2	0.037	0.574	0.612	0.175	0.212	0.387	0.341	0.081	0.423	0.755	0.005	0.760
4.280	9	4-8	3	0.048	0.604	0.652	0.228	0.224	0.452	0.434	0.085	0.519	0.854	0.005	0.859
4.510	9	4-8	3	0.048	0.634	0.682	0.228	0.238	0.466	0.434	0.089	0.524	0.854	0.005	0.859
4.970	10	5-9	4	0.056	0.658	0.715	0.275	0.248	0.524	0.512	0.092	0.605	0.912	0.005	0.918
5.200	10	5-9	4	0.056	0.682	0.739	0.275	0.261	0.536	0.512	0.096	0.609	0.912	0.005	0.918
5.670	11	6-10	5	0.063	0.703	0.767	0.316	0.270	0.587	0.576	0.099	0.675	0.946	0.005	0.952
5.900	11	6-10	5	0.063	0.723	0.787	0.316	0.281	0.597	0.576	0.102	0.678	0.946	0.005	0.952
6.360	12	7-11	6	0.069	0.739	0.809	0.352	0.289	0.641	0.629	0.104	0.733	0.966	0.005	0.972
6.590	12	7-11	6	0.069	0.755	0.825	0.352	0.298	0.651	0.629	0.106	0.736	0.966	0.005	0.972
7.050	13	8-12	7	0.074	0.768	0.843	0.383	0.305	0.689	0.673	0.108	0.781	0.978	0.005	0.983
7.280	13	8-12	7	0.074	0.782	0.856	0.383	0.313	0.697	0.673	0.110	0.783	0.978	0.005	0.983
7.750	14	9-13	8	0.078	0.793	0.871	0.410	0.319	0.730	0.708	0.111	0.820	0.985	0.005	0.990
7.970	14	9-13	8	0.078	0.803	0.882	0.410	0.325	0.736	0.708	0.113	0.822	0.985	0.005	0.990
8.440	15	10-14	9	0.081	0.812	0.894	0.434	0.331	0.765	0.738	0.114	0.852	0.989	0.005	0.994
8.670	15	10-14	9	0.081	0.821	0.903	0.434	0.337	0.771	0.738	0.116	0.854	0.989	0.005	0.994
9.130	16	11-15	10	0.084	0.829	0.913	0.455	0.341	0.796	0.762	0.116	0.879	0.991	0.005	0.996
9.830	16	12-15	11	0.086	0.854	0.940	0.472	0.363	0.836	0.782	0.122	0.904	0.992	0.005	0.998
10.300	16	16-15	15	0.128	0.871	1.000	0.615	0.384	1.000	0.871	0.128	1.000	0.994	0.005	1.000

TABLE 2

MIL-STD-781, TEST PLAN III

TEST PLAN IV

Test Parameters				True MTBF = $.5\theta_o$			True MTBF = $.75\theta_o$			True MTBF = θ_o			True MTBF = $2\theta_o$		
Time in Multiples of θ_o T	Number of Failures Causing Reject	Number of Failures to Continue Test	Number of Failures Allowing Accept	Prob. of Acceptance	Prob. of Rejection	Prob. of Decision	Prob. of Acceptance	Prob. of Rejection	Prob. of Decision	Prob. of Acceptance	Prob. of Rejection	Prob. of Decision	Prob. of Acceptance	Prob. of Rejection	Prob. of Decision
0.350	2	0-1	Not App.	0.000	0.155	0.155	0.000	0.079	0.079	0.000	0.048	0.048	0.000	0.013	0.013
1.040	3	0-2	Not App.	0.000	0.375	0.375	0.000	0.189	0.189	0.000	0.109	0.109	0.000	0.025	0.025
1.400	4	1-3	0	0.060	0.424	0.485	0.155	0.212	0.367	0.246	0.120	0.367	0.496	0.026	0.522
1.730	4	1-3	0	0.060	0.519	0.580	0.155	0.265	0.421	0.246	0.149	0.395	0.496	0.030	0.526
2.090	5	2-4	1	0.103	0.552	0.656	0.270	0.281	0.552	0.419	0.156	0.576	0.742	0.030	0.773
2.430	5	2-4	1	0.103	0.617	0.721	0.270	0.320	0.591	0.419	0.176	0.596	0.742	0.032	0.775
2.790	6	3-5	2	0.132	0.639	0.771	0.352	0.331	0.683	0.535	0.181	0.716	0.859	0.033	0.892
3.120	6	3-5	2	0.132	0.681	0.813	0.352	0.357	0.710	0.535	0.193	0.729	0.859	0.034	0.893
3.480	7	4-6	3	0.151	0.695	0.847	0.410	0.365	0.776	0.613	0.197	0.810	0.915	0.034	0.949
3.810	7	5-6	4	0.177	0.723	0.901	0.477	0.384	0.861	0.688	0.205	0.894	0.946	0.034	0.981
4.170	8	5-7	4	0.177	0.733	0.911	0.477	0.389	0.867	0.688	0.207	0.896	0.946	0.034	0.981
4.870	8	8-7	7	0.228	0.771	1.000	0.579	0.420	1.000	0.778	0.221	1.000	0.964	0.035	1.000

Table 3

MIL-STD-781B, Test Plan IV

RELIABILITY TECHNIQUES TO KEEP YOU OUT OF COURT

by

R. M. Jacobs and J. Mihalasky
Newark College of Engineering
Newark, N.J.

Introduction

The last several years has witnessed an astonishing increase in the annual number of product liability suits. Ten to fifteen years ago, the annual number of such suits was less than 5000. The current level is already in excess of 500,000 suits annually. The dollar value of the settlements of these suits has also risen dramatically. Six figure settlements are not uncommon; a few have even reached as high as eight figures in size.

Among the reasons advanced for the seriousness of this situation are three of interest to the Quality/Reliability Engineer:

1. Change in the legal attitude toward product liability from "Caveat Emptor" to "Caveat Vendor" - "let the buyer beware" to "let the seller beware";

2. The apparent decrease in the quality and reliability of product;

3. New legislation - local, state, and Federal - that attempts to protect the consumer of products and services.

Years ago, a manufacturer felt safe from product liability (PL) action because of the legal concept called "privity of contract" as outlined in the 1842 Winterbottom vs. Wright case. In this case the court stated in its opinion:

> "There is no privity of contract between these parties, and if the plaintiff can sue, every passenger, or even any person passing along the road, who was injured by the upsetting of the coach, might bring a similar action. Unless we confine the operation of such contracts as this to the parties who entered into them, the most absurd and outrageous consequences to which I can see no limit would ensure".

This essentially meant that the purchaser could only sue the persons with whom he had a contract covering the purchase of the product, i.e., the retailer. Since the purchaser did not obtain the product from the manufacturer, and the retailer had no part in the manufacturing of the defective product, the purchaser was left holding the bag. These were the days of "Caveat Emptor" -- let the buyer beware! This rule of privity, with few exceptions, remained well entrenched in the annals of jurisprudence of the United States until 1916.

In the 1916 MacPherson vs. Buick case, the buyer got a major break. In this case, MacPherson was driving a Buick automobile when the car collapsed. The New York Court of Annals, speaking through Justice Cardozo, held that the manufacturer was liable, in the absence of privity, for injuries resulting from the use of a product whether or not inherently dangerous if there was evidence of negligence in design, manufacture and assembly of the product. The court in MacPherson stated:

> "If the nature of a thing is such that it is reasonably certain to place life and limb in peril when negligently made, it is then a thing of danger. Its nature gives warning of the consequences to be expected. If, to the element of danger there is added knowledge that the thing will be used by persons other than the purchaser, and used without new tests, then, irrespective of contract, the manufacturer of this thing of danger is under a duty to make it carefully."

Thus the concept of privity of contract was abandoned, even destroyed. A product purchaser is now able to reach beyond his immediate contractual contract, in this case, the automobile dealer, and sue the manufacturer. It is important to note that in order to recover, the plaintiff had to prove that the manufacturer had been negligent. This requirement gave rise to a number of problems. These problems were partially solved by the theory of Warranty and the Uniform Commercial code. These instruments are not the subject of this paper.

Since that time, many additional legal decisions have opened wide the breach which allows a product consumer to sue "any and all" from the retailer through to the manufacturer, parts supplier, on down to the designer and quality engineer who may have contributed to the faulty product. The impetus for the most recent sequence of changes in liability law was derived from two significant cases, Henningsen vs. Bloomfield Motors and Greenman vs. Yuba Power Products Inc. The former was a case in which the plaintiff was injured, sued a dealer, the manufacturer of record and the supplier. The plaintiff was awarded a judgement which wiped out again the theory of privity and which also established the precedent that the manufacturer of record is responsible for the errors of his suppliers, even though the discovery of a defect by the manufacturer of record would have been difficult. In the latter case (Greenman) the purchaser of a power tool, a combination saw-drill lathe, sued the manufacturer. While the plaintiff was using the tool as a lathe for

turning a large piece of wood he wished to make into a chalice, the wood flew out of the machine and struck him on the forehead, inflicting serious injuries. The California Supreme Court held:

"A manufacturer is strictly liable in tort when an article he places on the market, knowing that it is to be used without inspection for defects, proves to have a defect that causes injury to a human being".

"The purpose of such liability is to insure that the costs of injuries resulting from defective products are borne by the manufacturer that put such products on the market rather than by the injured persons who are powerless to protect themselves."

These and other cases contributed to the development of the Restatement of Torts (Second) prepared by the American Law Institute. This body of law contained Section 402A in particular, which concisely summarized the recent products liability cases as follows:

<u>S402A - Special Liability of Seller of Product for Physical Harm to User or Consumer</u>

1. One who sells any product in a defective condition unreasonably dangerous to the user or consumer or to his property is subject to liability for physical harm thereby caused to the ultimate user or consumer, or to his property, if
   a. The seller is engaged in the business of selling such a product
   b. It is expected to and does reach the user or consumer without substantial change in the condition in which it is sold.
2. The rule stated in subsection 1 applies, although
   a. The seller has exercised all possible care in the preparation and sale of his product.
   b. The user or consumer has not bought the product from or entered into any contractural relation with the seller.

Essentially, this theory permitted those injured or suffering a property loss to sue, for financial satisfaction, anyone in the chain of commerce. This literally means any organization or anyone normally engaged in the sale of goods or services regardless of their relationship to those experiencing the loss.

Not only has this been a time of change in the law, the public attitude toward product quality and reliability has also changed. Mass production made most goods available, both in price and quantity, to the general public. But, the public was told that in return for mass produced, low priced goods, they had to be willing to accept some defective merchandise. These defectives were supposed to be an inherent characteristic of mass production. But as technology advanced, and products grew more complex, the price of these goods rose. The consumer began to be unwilling to accept the "you have to expect some defectives" theory for these new higher priced goods. With the improvement in communications, consumers began to publicize their problems and groups/agencies compared notes, and as a result the consumer became further dissatisfied with the acceptable quality levels tolerated by the manufacturer.

Product sophistication with its high price tag resulted in cost cutting competition. The cost cutting resulted in less expensive - often inferior - materials being used in the product in order to reduce its price.

This feeling of dissatisfaction by the consumer with product quality and reliability was sensed by public crusaders and politicians alike. City, state and the federal government enacted laws to protect the helpless consumer. Publicity was given to large product liability suit settlements, and crusaders such as Ralph Nader attracted a large following. The uproar was loud enough to cause the creation of a National Commission on Product Safety. Congress continues to discuss and enact more stringent consumer protection laws. It is expected that a Consumer Protection Agency will have been created prior to January, 1973. Most authorities believe that liability - safety - quality of product are related and Harry M. Philo, in a keynote address to the 1970 Product Liability Prevention Conference (PLP-70), stated: "Product liability will end only with product safety." Mr. Philo is author of the plaintiff attorney's bible "Lawyer's Desk Reference."

<u>The Reliability Engineer and Product Liability</u>

The reliability engineer is in an ideal position to serve as the key figure in any effort calling for the minimization of financial losses and the concurrent legal exposure due to a product liability event. There is no other "technical type" who normally uses, or has readily available to him, the techniques needed to minimize liability exposure. All that is needed on the part of the Reliability Engineer is a <u>change of attitude</u>. He has to think "reliable and safe", not just "reliable", as has been his custom to date. In this day and age it is not enough to have a "reliable" product. Many a reliable product has been unsafe, and has resulted in litigation against the manufacturers and distributors of that product.

Some of the standard reliability techniques and tools that are readily adaptable for product safety attainment are:

1. Reliability Prediction and Estimation
2. Failure Mode and Effects Analysis
3. Design Review
4. Human Factors and Maintainability
5. Maintenance and Failure Reporting
6. Subcontractor and Supplier Control
7. Standards Development

Reliability Prediction and Estimation.

It's function is to provide numerical estimates of the reliability of a system or of its subsystems. A product which "fails" in the terms of the Reliability Engineer could also fail and cause a tangible loss (injury, property damage or commerciall loss) Therefore, the frequency of failure is directly related to the frequency of a loss event subject to liability claims. When a product involving human safety has been sold and afterward dangerous defects in design have come to the manufacturer's attention, the manufacturer has a duty either to remedy the defect, or if complete remedy is not feasible, to give adequate warnings and instructions concerning methods for minimizing the dangers (Braniff Airways Inc. versus Curtiss Wright Corp.). The manufacturer also has a duty to warn the potential user or consumer when he knows that the use of the particular product in a certain way could create a danger. When he fails to give warning of such a known potential danger, a product sold without such a warning is in a defective condition, if it happens to be involved in an accident.

It is also clear that where the design of the product is changed so that it is not in the same condition as it was when it was manufactured or sold, and a loss occurs, recovery will be denied.

An unsafe or defective product frequently involves physical causes, with which a reliability analyst is familiar. He uses this knowledge in his prediction and estimation tasks, known as the physics of failure. Thus a tool familiar to the reliability analyst, prediction, can be used to estimate the frequency of the unsafe behavior of a product. He can examine the product's design and manufacturing methods, the quality assurance procedures that are used to detect and induce product flaws, and the interface between the product and its user. With this information in hand, he uses analysis, extrapolation, and combines them with other data to come up with a prediction as to the safe and reliable operation of the product. These predictions, when coupled with later confirming tests on off-the-shelf product, are potential elements of a manufacturer's defense in a court action.

The Reliability prediction and estimation can serve as a guide to a safer product by:

1. Evaluating the safety of one product design against another, and selecting the potentially safest design.

2. Determining the need for additional test information so that adequate safety information is available.

3. Evaluate the results of corrective design efforts initiated by production tests or field data.

Failure mode and effects analysis (FMEA)

This technique has as its purpose the minimization of failures, and hazards that affect reliable operation. The purpose can be extended to include in the definition of failures and hazards, a failure or malfunction that results in injury or death to a person damage to equipment, or commercial loss.

The analysis consists of reviewing each critical part of the design to establish what effect each failure or malfunction of this part will have on the safety of product's user. The result of this analysis may be the specification of a few new parts, or a major redesign. At the least the analysis should result in a fail-safe condition so that a failure will not result in a fire, explosion, or otherwise endanger life and limb. F.M.E.A.'s greatest potential in the reduction of liability exposure is its ability to uncover an unexpected weakness in a product that could or would result in an unsafe product.

Failure mode analyses should include the following:

a. Determination of the function of each part and sub-assembly.

b. Determination of all possible modes of failure.

c. Determination of the possible cause or causes of each mode.

d. Assessment of the effect of each mode of failure upon the product's performance and safety.

e. Estimation of the criticality or severity of the effect determined in (d).

f. Estimation of the probability of the occurrence of the particular failure mode. This may be quantitative if sufficient data is available, or may be categorized as high, medium, or low.

g. Recommendations for corrective action to eliminate the cause or reduce the probability.

Design Review

Design review, in its broadest sense, is a mechanism for complete review of all design data to assure that design features are such that the product is capable of being fabricated at lowest possible cost and still is capable of achieving the objective of successful performance under end-use conditions. No one man or particular design specialist can possibly know all of the ways to achieve this optimum compromise. Many types of consumer and industrial products are sufficiently complex, and operating requirements are sufficiently stringent, to warrant a "review-team" approach with each review team led by a responsible engineer.

The results of Design Review is synergistic in nature and serves as a method of communication. When these communications are dependent on informal meetings and memoranda, there are omissions. Some of the important organizations with whom early consultations must

take place are forgotten in the rush to get a new product started. The discussions that are held may be contrary to the most efficient and cost-effective way of planning the total program. For those reasons, it has been found that best results are obtained with the Formal Design Review, a planned and scheduled design review. Not only those directly concerned, but also other organizations with pertinent inputs are notified so that their contribution to the program can be offered to the development team.

Three design reviews are usually planned, because experience has indicated that as optimum. The first one is known as the Preliminary or Concept Design Review. This review is conducted to establish most of the ground rules and goals for the design. Some of the items considered and reviewed during the conceptual design review include:

1. Function to be performed by product.
2. Market and sales volume.
3. Design sequence (working elements and artistic appearance).
4. Subsystem concept (if applicable).
5. Make or buy consideration.
6. Subsystem interfaces.
7. Design Parameters (which are required in order of importance to function; some are mandatory, whereas others are only nice to have).
8. Test considerations.
9. Documentation required.
10. Critical parts to be used.
11. Environmental considerations.
12. High-risk areas (including product liability and safety problems).
13. Reliability requirements.
14. Redundancy requirements.
15. Schedule considerations and cost alternatives.
16. Establishing rank of importance for all requirements.

In this way, and at one meeting, all persons concerned with the design and program planning are a party to and are informed of the reasons for decisions made. In the event of a question that cannot be answered at the design review meeting, an "action item" is established and assigned to a specific person for study and a detailed recommendation. The design review is not considered complete until all action items are resolved. In this way, the chief design engineer -- or in a large program, the manager of the program -- keeps watch of who is doing what, when it will be completed, and how much it will cost.

The conceptual design review is followed at appropriate periods by the Interim Design Review, the Critical (or Final) Design Review and, if required, a Manufacturing Design Review. The number of design reviews and the timing of these reviews depends a great deal on the program. In small programs, one design review may be sufficient to satisfy all the questions involved, but in very large programs, many design reviews may be required.

As a part of, or as a supplement to these Design Reviews, there are related specific discussions on specifications, materials, parts, circuits, mechanical and electrical stress analyses and value analyses. These tasks consider the -

1. elements of describing the product for those being purchased or obtained from other profit centers,
2. provisions for reduction of the adverse effects of thermal, chemical, radiation, vibration, and shock environments,
3. adequacy of materials employed,
4. reduction of human error,
5. maintenance provisions,
6. production cost reduction (value engineering),
7. estimated failure rates,
8. estimated repair or maintenance rates, and
9. failure mode analysis.

Human factors and maintainability

One of the major areas for potentially unsafe product performance is the product-person interface. The way to reduce the potentialities in this area is to give careful consideration to the elimination of human-induced error. Particular attention is paid to the areas of serviceability, maintainability, and installation in-so-far as the product and the human is concerned.

However, the human element of anticipating how the product might be used is most difficult and vitally important. In attempting to foresee potential uses the application of the synergistic Formal Design Review is most useful. The importance to "foresee" the use of a product is both technical and legal. Technically, a decision can be made to design for that application or to design the product to forestall that particular use. If the use cannot be prevented and is a potentially dangerous application, then a warning may be required to alert the user to the risks involved.

Legally, the effort applied to anticipate potential applications is only valuable in the event of a liability suit. It can then be shown that the designer/manufacturer and/or others were not negligent in performing their tasks and attempted to "foresee"

potential uses. Even when an application was not anticipated but caused a loss, the award/settlement was substantially lower than requested since the effort was made and could be proven. The assurance approach to the design and operational characteristics of the product are analyzed and are also evaluated against human factor and load considerations.

The basic approach to this is:

1. Break down the operation of the product or service to be performed, into functions
2. Select hardware approaches to perform each function, deciding at the same time which of the functional will be manual, and which automated.
3. Establish basic installation, servicing and maintenance approaches.
4. Continue the approach from the initial design stage, through the preproduction stage. As the product design progresses through to the production stage, bring the human factor considerations into sharper focus.

The result of honestly using this sequence has to be simplification of installation, serviceability and maintenance activities; accessibility for servicing and maintenance, and clear and effective procedures that result in error free activities. There also has to be a minimum of stress and confusion to the operator, to the maintainer, and to the installer.

<u>Maintenance and failure reporting and Correction.</u>

The reliability engineer has long relied on field data on maintenance problems and failures. Data on product failures from service personnel, from test facilities, and from test laboratories is also a valuable technique for minimizing liability exposure. An efficient reporting system can result in product correction before large quantities of product get out into the stream of commerce, or in a product recall before there has been a major exposure by the public to an unsafe product.

A data feedback system must satisfy internal organizational requirements as well as technical requirements of the product. For such a data system to achieve the necessary results, it must incorporate the following essential features:

1. A procedure for identifying each end product and its constituent or field replaceable parts.
2. Establishment of consistent, confidential and multiple data sources and input information.
3. A method of recording and reporting on the length of "acceptable" product operation.
4. An easy method of recording and reporting on the various details pertaining to product malfunction.
5. A method of assuring implementation of proper and timely corrective and preventive action, based on efficient data retrieval methods.
6. An effective information feedback system to insure that all parties receive timely and accurate data (including customers and suppliers).

The type of questions that the system (specifically the product user or serviceman) can supply answers to are:

1. How long does the product operate satisfactorily?
2. How often does the product fail?
3. Which item(s) in the product cause(s) failure?
4. Do these failures endanger lives or property? Could they cause harm?
5. Do failures occur within the warranty period?
6. How long after the warranty period expires do the failures occur?
7. How much do the failures cost the manufacturer, a) while the product is in warranty; b) after the product is out of warranty as a "policy" fix or as a marketing fix?
8. What does it cost the customer in time and money to repair the product?

Of course, the President is always asking - Have we lost the customer? Could we? There are very few opportunities to answer these questions without many qualifications and thus are unanswerable.

<u>Subcontractor and Supplier Control</u>

Because of the use by modern industry of a system of subcontracting to acquire parts and subsystems, the reliability engineer has had to establish a method to assure the reliability of these purchased items. This approach works equally well in assuring that purchased parts and subsystems will result in a safe product. The key to this control procedure is to insure that the subcontractor and supplier is taking the same precautions to insure a safe product that the prime fabricator takes. The essential features of an efficient program are:

1. Selection of vendors and subcontractors who have demonstrated their capability to produce a safe and reliable product.
2. Development of adequate specifications and test procedures for the purchased items.

3. Development of proper safety, reliability, and quality program requirements to impose on the subcontractor.

4. Establishment and maintenance of effective communications with the subcontractor in order to minimize misunderstandings, and facilitate identification and correction of problem areas.

5. Continuing audits to insure that the subcontractor is implementing the agreed upon reliability, quality and safety program.

Standards Development

The practice of having industry, company and division standards was born to 1) introduce the costs of redesigning "the wheel" each time it was needed; 2) to minimize the cost of manufacture; 3) to permit interchangeable replacements for purchasers of product, and 4) to provide a set of prearranged requirements for a product so that it could be purchased on a competitive basis from many sources and to permit simplified communication between buyer and seller.

Thus Standards have been developed to describe a product's function and/or its manufacturing process. In both types of standards, physical dimensions, and, where applicable, material and electrical characteristics, are usually specified.

This practice has developed in such a fashion that a "standard" has been the description for the item that just satisfies the requirement.

As a result, quality, reliability and safety requirements are minimum requirements for the product. Although Standards are relied upon for purchasing, manufacturing and descriptive purposes, caution should be exercised in the area of safety, as few Standards quantify the requirement. In general, a Standard is not admissible in Courts as a defense unless the author is available to defend its adequacy. Non-compliance to a standard is admissible by the plaintiff to show inadequacy, even to the minimum requirements.

The use of standards by the Reliability Engineer becomes an excellent starting point. The specialist should expand the Standard with addendum requirements, including testing, safety, life of operation, reliability and other characteristics which can be quantified or otherwise measured. If evidence is available that denotes a conscious and deliberate effort to provide a safer, more reliable product by use of basic standards, plus these additions, the inclination will be to consider this information heavily in legal deliberations. However, it must be noted that if a product does not satisfy the minimum requirements of a standard, it will be a very difficult task to successfully defend a liability action involving the product.

Therefore, standards are to be considered an absolute minimum to which requirements and tests should be added to satisfy peculiar application requirements.

Conclusion

The plaintiff does not automatically win every claim filed. Although, according to Jury Verdict Research, Inc., plaintiffs have won in an increasing number of cases. See Table I for details. He is still obligated to prove five things:

1. That the defendant is engaged in the business of either manufacturing, selling, distributing, or supplying the product, or engaged in the business of renting or leasing such product.

2. That the product contained a condition that was unreasonably dangerous.

3. That the condition existed at the time it left the defendant's control.

4. That the plaintiff sustained injury.

5. That the unreasonably dangerous condition was a proximate cause of the injury.

The Quality/Reliability specialist is involved directly with the known defenses - the plaintiff assumed the risk (warnings and prior knowledge) - the plaintiff grossly misused the product well beyond all anticipated or even surmised applications (foreseeable applications - the plaintiff caused the failure by his own actions, or lack of.

The point has to again be made, and stressed, that the reliability engineer already is familiar with, tools and techniques that are needed to minimize a manufacturer's product liability exposure. These reliability tools and techniques exist, and all that is needed is for the reliability engineer to change the emphasis of his function from "aiding in the design and manufacture of a reliable product", to that of "aiding in the design and manufacture of a safe and reliable product".

The Reliability Specialist should become especially acquainted with the law of torts in the States where his product is sold and used. In particular he should become acquainted with the Court of Appeals decisions involving all products. He should not confine his research to his own product areas. Law is an ever evolving specialty and requires continued surveillance and translation for the designer and manufacturing personnel. Above all, one or two trial court cases which prove a point does not necessarily establish the law. State and Federal Appeals Court findings are used to establish precedential law and are usually the referenced material in court briefs.

The courts have frequently taken the position that anyone who enters a special field of manufacturing will be held to possess the knowledge and skill of an expert in that

field and must keep reasonably abreast of techniques and devices used by practical men in their trade. As a result, the manufacturer must avail himself of the expert and specialized knowledge which may exist as to proper and reasonably safe design of the particular product involved, and he cannot close his eyes to what is known to other experts.

Industry	1960-1966	1966-1971
Drug-Pharmaceutical	56%	72%
Industrial Equipment	48%	55%
Automotive/Truck	32%	47%

Percent of Liability Suits Won by Plaintiff
Source - Jury Verdict Research, Inc.

TABLE 1

General References

1. National Commission on Product Safety, Final Report, June 1970.

2. National Aeronautics & Space Adm'n, <u>An Introduction To The Evaluation Of Reliability Programs</u>, NASA SP-6501, Office of Technology Utilization, NASA, Washington D.C., 1967

3. Hulme, H.D., "Data Collection Systems - A Tool for Product Liability", <u>Product Liability Prevention '70 Conference</u> Proceedings, Newark, N.J., August 1970, pgs. 15-19.

4. Hulme, H.D., "Products Liability and Product Manufacturing", <u>Annual Technical Conference Transactions</u>, A.S.Q.C. Chicago, Ill., May 1971, pgs. 381-386.

5. Kerekes, Emery B., "Product Liability and the Manufacturing Manager", <u>Proceedings Product Liability Prevention Conference</u>, Newark, N.J., August 1971, pg. 131.

6. Jacobs, Neville E., "Prediction as a Prevention Tool in Product Liability", <u>Proceedings of PLP-72</u>, Newark, N.J., August 1972.

7. Jacobs, R.M., "Implementing Formal Design Review", <u>Industrial Quality Control</u>, Vol. 23, No. 8, February 1967, pgs. 398-404.

8. Jacobs, R.M., "Minimizing Hazards in Design", <u>Trial Lawyer's Quarterly</u>, Vol.7 No. 2, Summer/Fall 1970.

9. Jacobs, R.M., Okun, A.M., and Dertinger, E.F., <u>Pacer</u>, Vol.1, No.4, July-Aug. 1969.

10. Jacobs, R.M., and Mihalasky, J., "Product Liability - A Challenge for the I.E.", <u>Industrial Engineer</u> the Journal of the AIIE, November 1972.

11. Jacobs, Wyatt, "The Design of the Growth of Product Liability Suits to the Engineer", A.S.M.E. Paper 72 DE 14.

12. Mihalasky, J., "The Cost of Unreliability in the Market Place", <u>Proceedings of IP-70-IEEE</u>, October 1970.

13. Mihalasky, J., "The Product Liability Program - Its Effect on Industry", <u>Transactions of the 23rd Annual Conference</u>, Metropolitan Section, American Society for Quality Control, Rutgers University, Sept. 11, 1971, Pgs. 74-78

14. Mihalasky, J., "The Great Litigation Confrontation - Its Effect on Industry", <u>S.A.M. Advanced Management Journal</u>, Vol. 27, No. 1, January 1972, pgs. 56-60.

15. Schirmer, Walter E., "The View from the President's Office", <u>Products Liability and Reliability, Some Management Considerations</u>, Machinery and Allied Products Institute, 1967.

16. Schroeder, R.E., "Consider the User", <u>Product Liability Prevention '70 Conference Proceedings</u>, Newark, N.J., August 1970, pgs. 149-155.

17. Smith, Marion P., "Design Review as a Tool for Product Liability Prevention", <u>Proceedings of PLP-72</u>, Newark, N.J., August 1972.

18. Smith, T.H.F., "Documentation for Liability Prevention", <u>Product Liability Prevention '70 Conference Proceedings</u>, Newark, N.J., August 1970, pgs. 111-115.

19. Weiss, Stanley and Baker, G. "Materials Failure as a Cause of Product Liability" <u>Proceedings of PLP-72</u> August 1972, Newark College of Engineering.

20. Wells, Robert L., "Prevention of Product Liability Hazards in Manufacturing", <u>Products Liability and Reliability, Some Management Considerations</u>, Machinery and Allied Products Institute, 1967, pg. 23.

21. Woods, G.W., "Total Organizational Legal Risk Reduction Considerations", <u>Annual Technical Conference Transactions</u>, A.S.Q.C., Chicago, Ill., May 1971, pgs. 367-372.

22. "More Plaintiffs Win Product Liability Suits" <u>Business Insurance</u>, August 14, 1972.

Legal References

1. Winterbottom vs. Wright, 152 Eng. Rep. 402 (1842)
2. MacPherson vs. Buick Motor Co., 217 N.Y. 382; 111 N.E. 2d 1050 (1916)
3. Henningsen vs. Bloomfield Motors Inc., 32 N.J. 358; 161 A 2nd 69 (1960)
4. Greenman vs. Yuba Power Products Inc., 59 Cal. 2d 57; 377 P2d 897; 13 A.L.R. 3d 1049
5. Braniff vs. Curtis Wright, 411F.2d 451.
6. American Law Institute, Restatement, Second, Torts - Section 402A (1962)
7. Harper & James, Section 28.

ENGINEER'S LIABILITY IN PRODUCTS LIABILITY

Fredric D. Fogle
Q.V.S. Inc.
East Orange, New Jersey

The engineer's potential personal liability in product liability suits has never been greater than today, yet the subject of personal liability for product related losses has not received the attention showered on the headline cases involving strict product liability. Some articles have minimized the existence of personal exposure, creating an unwarranted sense of security. The risk of personal liability for losses caused by products with which the engineer works is real, and is increasing at least in proportion to the general increase in product liability litigation.

Strict Product Liability

The bulk of attention since 1960 has centered on the evolving field of strict product liability, that is - liability without regard to fault. This burden is placed upon the manufacturer as the one best able to allocate and trade off the costs of improvement, inspection, testing, insurance or risk of loss over the broadest base, that of consumer sales. In many instances this same logic applies to distributors, retailers, installers, servicemen, etc.

Strict product liability has now reached the point where the manufacturer and others in the chain of distribution are liable to anyone suffering a loss of any kind attributable to the use of an unsafe or defective product. The usual vehicle to accomplish this is strict liability in tort, which eliminates all traditional privity and fault requirements, with recovery being subject only to minimal care by the user in application, use, and maintenance. Thus a user may not recover if he has elected to use a known defective product, or if he uses an otherwise safe product in an unreasonable manner.

The burden of proof that a given product is unsafe or defective and caused the loss still rests upon the injured party. If the manufacturer can show that the condition did not exist at manufacture or as a result of manufacture, he will avoid liability. Efforts spent to reduce the incidence or effects of failure will of course directly reduce the total liability exposure, but strict liability is not a form of presumed negligence and all the data in the world such as inspection standards, training, testing, reports, etc., will not help the defendant at trial unless it also proves that the specific product involved was not defective.

A product can be defective as a result of inadequate warnings, markings, instructions, or safety devices, although otherwise adequate. This becomes a considerable burden since any reasonably forseeable use must be allowed for. Use would include storage, transit, application, and disposal. Even reasonable modifications made by others may not preclude liability. The liability of component manufacturers for end product defects traceable to their components is unsettled.

Strict product liability can apply even if it is beyond the finest state of the art to detect or correct the defect, and is an allocation of risk somewhat like insurance. A difficult legal problem exists where strict liability is applied to situations where the consumer base is small, the hazard of use is large, but the social benefit is immense, as in medical implant devices. The added financial burden of potential loss may make the product cost prohibitive to the user or result in the withdrawal of the product by the manufacturer, neither result being acceptable.

The source of strict product liability is judicial, not statutory, and your actions of today will be judged by standards set tomorrow. Thus a power press made in 1949 can be judged defective for inadequate safety devices by 1972 standards based upon a cause of action which didn't even exist until 1960.[1] In another case a charcoal manufacturer was held liable for failing to warn of possible carbon monoxide poisoning when the user burned the briquets indoors without ventilation.[2]

Enterprise Liability

The difficulties of proof that a product was defective or unsafe, and its causative relation to the injury, have created two trends of thought. The first simplifies the burden of proof to simply showing the likelihood that a defect caused the loss, and is based upon a presumption that safe products don't fail.[3] The second persues the elimination of strict product liability as it is now, and its replacement by "Enterprise Liability", a term sometimes also associated with the reduced burden of proof proposals. Enterprise liability is a proposed form of absolute liability for all injuries involving products, regardless of failure or defect. It is somewhat analagous to present workman's compensation statutes and seeks only to determine that a loss occured as a result of the use of a product or group of products. It responds to the philosophy of those who believe that social goals should be base-loaded on business, and also of those who forsee undesired direct government control as the only other alternative.

Clearly society should seek a basis for assigning risk which rewards care, penalizes malice or indifference, and allocates incidental losses over a large basis when possible. Enterprise liability does only the latter and will probably not replace strict product liability, with possible exceptions in areas such as children's products, unless strict liability fails to meet the social demands placed upon it.

Liability of the Engineer

There has been nothing in the above philosophies of strict product liability or enterprise laibility which would lead to assessing personal liability on individual engineers; but the overall field of product liability is much larger and includes other forms of legal action, such as:

- Uniform Commercial Code (UCC) express and implied warranties
- Strict warranty laibility
- Breach of contract
- Criminal proceedings
- Intentional torts
- Negligence

An action begun as a strict liability claim against the installer may come home to the engineer as an action for negligence, etc. These were the means for many recoveries before strict liability in tort and are still available to injured parties. In states which have not adopted strict liability in tort a similar result is obtained by use of strict warranty liability and the associated UCC provisions, with difficulty primarily arising over statutes of limitation which run from the date of sale. Since interest here is in the engineer's liability, it is noted that the first three actions listed are effective against sellers, which normally excludes engineers, although caution is needed by those who consult, are registered professional engineers, are partners in engineering concerns, or own or work for unincorporated businesses. This leaves at least three areas for potential individual liability.

Criminal Liability

Criminal liability is meant to protect the public interest by punishing, physically or monetarily, for a violation. It seeks to deter others from such behavior and therefore is appropriate primarily where voluntary acts have caused an injury. State laws vary widely but normally provide for criminal penalties whenever death occurs from willful, reckless, or negligent behavior and whenever bodily injury results from willful or reckless behavior. Likely cases for criminal action would be those where the obvious result of the engineer's action could be bodily harm, such as - approving shipment of known spoiled food, utilizing defective safety devices, covering up product failures, etc. It makes no difference whether these acts were unilateral or at the direction of others, unless in fact a superior has acted in place of the engineer. Such a circumstance is difficult to prove without written evidence and a jury could just as likely conclude that both acted in concert.

Tort Law

Tort law provides monetary compensation to the injured party. It allows recovery for acts subject to criminal prosecution in addition to the criminal penalty, and also provides recovery where negligent behavior has resulted in physical injury or monetary loss. The classification of behavior as intentional, negligent, or normal is more difficult than meets the eye. The same result may obtain where a person intentionally does a harmful act, where he permits it to happen through negligence, or where it occurs out of his necessity to avoid a greater harm. The gray area exists where an action likely to cause harm is intentionally commenced but without intent to cause harm.

Personal liability for intentional torts is a reality in product liability. There is little doubt that an employee can be held personally liable for the consequences of intentional acts such as the knowing removal of safety devices, passing known deffective material through inspection, falsifying data or records which mislead others, or covering up of ones own mistakes or those of his subordinates. Punitive damages are appropriate for intentional torts.

Negligence

The biggest portion of tort law applies to loss from negligence. Negligence as a vehicle for recovery of damages requires more than its name implies. Basically negligence consists of an unreasonable violation of an obligation of conduct, which becomes a material factor and substantial factor in causing harm related to the conduct. The work of production employees and some engineers is so subject to further supervision and inspection to logically preclude individual liability since the employee has every reason to believe that any errors he has made will be caught before shipment, and in fact may not even be expected to note or mark discrepant material. Most engineers, however, have actual decision control over some aspect of a product that they know controls its ultimate safety. It is not sufficient defense that the company president can always order changes made. Certainly any engineer who creates a hazardous product through negligent performance of his duties can be liable for the consequences. Using the same examples as under intentional torts, he may have removed safety devices for overload test and failed to replace them, he may have erroneously labelled one item as another, or failed to review a report showing defective material about to be shipped.

The real problems facing the dedicated engineer are very difficult. He will be liable when reasonable behavior on his part would probably have prevented or helped prevent the defect from causing injury. When product liability was limited to the intended or normal uses of the product it was possible to balance economic and performance requirements to obtain a suitable product design. Under strict liability in tort all reasonably forseeable uses must be considered, even those involving improper maintenance, unusual environments, etc. If a concern does not make the end product the problem becomes overpowering. A manufacturer of bolts cannot envision all the reasonable uses of his product, and the pricing requirements for uncritical applications precludes use of only stainless steel or 100% testing for defectives. Is it not reasonable to expect the end item designer to allow for a bolt

pattern which retains adequate strength even with failure of an individual bolt, or to state specifically any purchase requirement where product safety is involved or which exceeds normal commercial usage? The law has not resolved these problems because it does not understand them. An engineer realizes that a pressure controller will fail someday, and designs a relief valve or backup into the product. Most states today hold the manufacturer of a component responsible for end item failure even where the end item manufacturer has ignored suggested usage precautions. New York notably has held for the component manufacturer but not for reasons related to component suppliers in general.[4] It would seem appropriate to place a burden on the end item manufacturer to adequately allow for backup systems and protective measures or bear the risk of loss from reasonably expected failures. If the end item manufacturer does not provide adequate design, the subassembly and component manufacturers will each allow redundant design safety factors and the safety dollar will be wasted.

Older cases limited the employee's duty in tort to the non-negligent performance of duties required by his employer. With the development of awareness of product liability it is evident that the reasonable engineer will be expected to think of the user and public interest as well. The Canons of Ethics for Engineers clearly include these duties.[5] Is it negligence to design an aerosol can without a safe means of failure if inadvertently disposed of in a trash fire? The answer will come from a jury. It is apparent that such situations can arise in design, testing, fabrication, assembly, sales, quality control, administration, packaging, servicing, or even advertising. The small company engineer may face several facets of liability at once.

Engineers must consider the modes and consequences of failure and accident as well as their prevention. As expectations of the public change, the standards of reasonable engineering performance change with it. Fortunately, product law does not yet require the engineer to play "Ralph Nader" by asserting himself into decision areas assigned to others. Product liability extends to circumstances about which the plaintiff may simply lie, the number of bottles which explode when picked up stretch ones belief.

Deep Pocket Theory

It is often stated that engineers will avoid personal liability because injured parties will prefer to bring suit against the wealthiest party, usually the manufacturer, on the easiest grounds of strict liability. This may be true in many cases but is little consolation to the engineer who is brought in for personal liability. The injured party will look at all the available defendants and the causes of action against each, compare this with their wealth, and will proceed on basis which offers the best potential awards and settlements. Since adding claims and defendants costs very little until trial, a typical suit will also include claims based on negligence and warranty against whomever in the chain of distribution or manufacture is available for service of process. Each defendant will file cross-claims and counterclaims against parties already involved and third party claims against previously ommitted parties, seeking to have others assume or at least share any liability.

Engineer as a Defendant

Someplace about here is where someone may think of the engineer, certainly his employer and its insurer are aware of him. If not, the parties proceed with "discovery", a legal tool whereby others can review the records of the product from conception to date, interview all levels of employees, and even question the employer's expert about many matters.

By now the chances are better that someone has thought of the engineer, particularly if his name keeps appearing on product documents. It should be noted that there is no "5th amendment" in civil suits and that statements made by parties to the suit are an exception to the heresay exclusion. Unfortunately it is also true that the case will be tried by attorneys who may be poly-sci majors, in front of judges who may be politicians, and jurors who won't understand anything said in engineering terms.

Many factors work in favor of the engineer, either by his omission from the suit or by subsequent events, such as:

A- Statutes of limitation, and the date from which they run, vary from state to state and according to the type of action persued.

B- It may not be convenient to try the case in a state where personal jurisdiction over the engineer is available, normally his states of residence and employ.

C- The proofs needed for intentional and negligent torts are more involved than for strict liability.

D- Unless the engineer has obvious wealth, a judgement against him may be worthless. Normal homeowners insurance provides no coverage and thus does not induce suit.

E- The engineer is presumed to have only that knowledge and ability necessary to perform his duties.

F- Individuals receive favorable consideration from the jury.

G- The employer is also liable for the negligence of his employees without proof of even which employee was negligent, and under strict liability is liable for anything which causes defective products to reach and harm the consumer.

H- The employer must indemnify the employee for any loss attributable to specific employer demands or instructions.

Other factors work unfavorably, such as:

1- Other defendants may no longer exist due to merger, sale, dissolution, or bankruptcy.

2- Other defendants may be in poorer financial condition, either now or at judgement date.

3- Having the engineer as a defendant may allow suit where conditions favor the plaintiff.

4- Punitive damages available for intentional torts may increase an otherwise nominal loss.

5- It does not cost much to include the engineer in a suit.

6- Having the engineer as a defendant makes him less believable as a witness for the employer.

7- With the emphasis on product documentation there is an excellent chance the engineer has already signed documents displaying his negligence, or has at least shown himself as being responsible for actions which may have caused the defect.

8- Violation of a statute implies negligence, but conformance does not preclude it. Statutes are a minimum standard of care, including OSHA.

9- It is well established that an otherwise innocent employer may be indemnified for any loss caused by the negligence of his employees. The insurance company may likewise recover for losses it has paid for the company.

10- Other employees may be brought in personally and decide their supervisor is the best one to blame everything on.

11- If the engineer has changed jobs his ex-employer has little remaining motivation to protect the engineer, and has all the records he left behind.

12- Product liability suits usually involve a seriously injured plaintiff who captures the sympathy of the jury.

13- If employers respond to product liability with insurance instead of safe products, social and legal pressure on the engineer for liability will increase.

What to do? There are at least two means of avoiding individual liability, or reducing it. One is to make the product safer to eliminate all liability. The other is to reduce the liability of the engineer for whatever losses do occur.

Volumes have been written about how to reduce failure through product control using techniques such as design review, testing, sampling, recall, feedback, projections, etc. The engineer will still usually be calling for less than he knows is optimum, for economic reasons. Regardless of the evolution of strict liability the engineer remains liable for negligence. If awards become common against engineers a trend away from the unique and progressive towards the usual and mediocre will result.

Action to Reduce Engineer Liability

Probably the most obvious solution is to cover engineers on the same insurance policy protecting the manufacturer, although the extra cost may be out of proportion to the risk. The benefits of having the same insurer are apparent since the engineer and employer need to cooperate to present the best defense and can only cooperate if adverse interests are minimal. Individual coverage on an individual policy is also feasible but probably not practicable at the present time, unless some group were to broadly sponsor such coverage.

Other action is available where direct coverage is not obtained. Until the questions surrounding employer/insurer indemnification from the engineer are resolved the following precautions should be considered:

I- A hold-harmless clause in the employment contract would preclude indemnification to the employer/insurer and could provide indemnification to the engineer for any liability to others.

II- A covenant against suit would preclude liability to the employer/insurer but not to others. Such a covenant could also be part of the employment contract.

III- A release should be obtained before cooperating with the employer in trial preparation, particularly if the engineer is a defendant or can still be added as a party. Releases are also appropriate against any party who wishes more cooperation than is granted by the discovery rules.

Any of these means could be modified to cover only negligence if the employer does not wish to become laible for intentional acts. The employer's insurer may not wish to be bound by any agreements, in which case the engineer must feel confident that the employer can actually hold him harmless without aid of insurance.

If no protection can be obtained from the employer or its insurer, the engineer must consider tham potential legal opponents, not a healthy situation. Engineers are used to signing documents that keep business moving and are interested in seeing the employer stay healthy, but do not look forward to having their own reports used against them by their employer or others. Most engineering problems are problems because of the need for critical decision among various solutions, a perfect target for the hindsight expert.

Many records are kept for initial product design, evaluation, liability review, and quality control and are then needlessly retained after all use is past. While one hesitates to advocate destroying records which may link the engineer to a specific

product loss, this may be the only record retained linking anyone specifically to the product design. There is certainly no responsibility to retain records just to help others in a lawsuit.

Legislative Action

None of the reasons for advancing the scope of strict product liability or of enterprise liability are helped by allowing individual engineer liability as well. The concepts yielding individual liability are based on old common law concepts no better established than those which have moved aside for strict product liability. The judicial allowance of individual liability is still unclear, but can be met with legislative action which would abolish individual liability for negligence which results in product defects where another source of recovery is available and is appropriate. The homeowners coverage could also be extended to include employee losses relating to products claims under the general liability clause, perhaps as a rider. Enterprise liability is not a reality yet and will probably only come about as a result of legislation. Engineering organizations should press for specific exclusion of derivative employee liability when and if such legislation is considered.

References

1. Bexiga v. Havir Manufacturing Corp., 60 N.J. 402 (1972)
2. Alferi v. Cabot Corp., NY AppDiv 2d 455, 235 NY 2d 753
3. Heaton v Ford Motor Company, 248 Or 467, 435 P 2d 806
4. Goldberg v. Kollsman Instrument Corp., 12 NY 2d 432, 191 NE2d 81
5. Canons of Ethics for Engineers, ASME bylaws, Novemebr 1948
6. Uniform Commercial Code, American Law Institute, 1962
7. Prosser on Torts, West Publishing Co.
8. Harper and James, Torts
9. For further reference:

   Henningsen v. Bloomfield Motors Inc. 32 NJ 358, 161 A2d 69

   Greenman v. Yuba Power Products Inc. 59 Cal2d 67, 377 P2d 897

   Vandermark v. Ford Motor Co. 61 Cal2d 256, 391 P2d 168

   Santor v A&M Karagheusian, Inc. 44 NJ 52, 207 A2d 305

   Burbage v. Boiler Engineering & Supply Co. 433 Pa 319, 249 A2d 563

RELIABILITY TECHNIQUES A N D PRODUCTS LOSS CONTROL:

Russell E. Marhefka
Consultant - Products Liability
Insurance Company of North America
Philadelphia, Pennsylvania

ABSTRACT

The major purpose of this paper will be to place the role of reliability technology in its proper perspective with regard to the overall concept of Products Loss Control. I intend to define, from the insurance industry point of view, the Products Liability situation as it exists today and the programs the insurance industry has been required to initiate in both the Technical Service and Underwriting areas in order to attenuate the Products Liability problem. In addition, one of the major program areas, the development of Product Loss Control Consultation capability, will be expanded upon. The expansion of this subject material will include a discussion of those reliability techniques, associated with a design review (FMEA, Fault Tree Analysis, prediction techniques, testing, etc.), that are currently being reviewed and evaluated by insurance industry engineers as one facet of their growing Technical Service capabilities.

CURRENT PRODUCTS LIABILITY ENVIRONMENT

An insurance company, like most businesses, is extremely sensitive to financial trends and results. INA as well as the rest of the insurance industry has lost money in product liability coverage in increasing amounts since 1967. These losses reached a record amount in the millions of dollars in statutory underwriting losses during 1971. These products liability losses have been precipitated by the current social, legal, and economic environmental aspects of the problem.

In the past six or seven years not only has the ballpark changed but the rules of the product liability game are different, as a result of:

1) A growing public awareness of legal rights due to:

   a) Publicity
   b) Legislation
   c) Plaintiff's Attorneys Associations

2) Loss of Privity

3) Evolvement of "Strict Liability" concept

4) Increased awards

5) Inflation

The preceding factors have resulted in a deteriorating Products Liability/General Liability loss ratio. In 1969, losses in Product Liability represented 34% of total General Liability losses. In 1970 and 1971 this figure had become approximately 50%.

The average number of Product Liability cases/week, being received through 1971, approximately tripled as compared to 1970. More important, these potential losses emanate from many different product liability loss sources, including:

1) Bodily Injury

2) Property Damage

3) Business Interruption

4) Extra Expense

5) Loss of Income

and are not limited to the more common categories ((1) and (2)) as in the past.

INSURANCE INDUSTRY REACTION

A. UNDERWRITING RISK EVALUATION

The impact of the burgeoning Products Liability problem has precipitated the realization by the insurance underwriter that information that can be developed for him by his company's Technical Personnel first; forms a solid base which allows him to realistically rate and competitively price a risk so that his company will keep their good accounts and, secondly, enables the underwriter to intelligently write new business in the difficult Product Liability market.

In order to accomplish these two goals it has been necessary for the underwriter to formulate the following formal outline for the thought processes that should be utilized when he is attempting to evaluate a risk for acceptance or declination purposes.

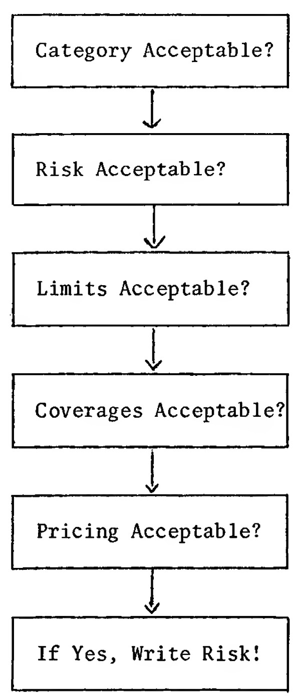


Figure 1. Risk Evaluation Format

1) Unsafe product design

2) Inadequate manufacturing and quality control procedures

3) Inadequate product hazard warnings and instructions

4) Misleading representation of products

Up to now the manufacturer has been able to live with his inadequate design review, reliability, quality control, testing, manufacturing, and production techniques, if he so desired. This was primarily because the manufacturer didn't have to overly concern himself with the consequences.

However, the law and the social environment is changing, and as a result the number of product liability suits in the courts are skyrocketing as are the associated awards. In addition, some of the most astute legal minds on the bench have stated that the cost of products liability should be imposed on the manufacturer regardless of fault since he is best situated to distribute the total cost to all users of the product. And consequently, the manufacturer should be advised that the insurance companies are no longer going to needlessly bear the brunt of the passed-on costs of industry's lack of commitment to product quality and safety.

In the past, the manufacturer has, generally speaking, decided the economic level at which he will cease to invest money into product safety and invest, instead, either in purchasing insurance or in the cost of meeting settlement demands in product liability cases. Thus, the manufacturer probably arrives at his optimal investment, but his criterion is totally unacceptable to the consumer who gets the defective product and who may consequently meet with a serious accident as a result.

Quality and safety have always been carefully considered, but mainly because poor quality or a poor safety record would be bad for business. Both quality and safety have been frequently sacrificed in order to make a product saleable. In other words, to keep the cost low.

Being a reasonably pragmatic individual I feel that tremendous economic pressure is now going to be placed upon the manufacturer, by the insurance industry, in the area of products liability insurance coverage. This pressure will be created by the insurance industry's desire to impress the manufacturer with the economic importance of a significant Products Loss Control Program effort and will be implemented through the substantial enhancement of the insurance industry's technical service and risk review capabilities.

In short, the insurance industry wants their clients to knuckle down to a rigorous approach to assure high standards of safety and quality control and are training consultants/specialists to evaluate their clients ability to do so.

E. MAJOR DEPARTMENTS AND ACTIVITIES INVOLVED

The major departments and activities that must be incorporated into a satisfactory Products Loss Control Program include:

I. Manufacturing

II. Quality Control

III. Marketing

IV. Record Keeping

V. Complaint/Incident/Accident reporting and investigation

VI. Product Design

I would normally discuss Product Design first, because I consider this department or activity to be the most important or critical in the development of a safe, reliable product. However, today I will discuss it last and attempt to expand or elaborate somewhat on the extremely significant role that reliability technology plays in the ultimate design prevention of serious products liability exposures.

I. PRODUCT MANUFACTURE

With regard to product manufacture, the insured's:

a) Production facilities must be adequate

b) Employees must be skilled, stable, and have pride in their work and product.

Associated with product design and manufacture is an extremely important, and many times overlooked facet of a products loss control program; the determination and evaluation of past or discontinued products or product lines. Discontinued products or product lines can be the source of very expensive exposures.

II. QUALITY CONTROL

A comprehensive Quality Control Program is a must in order to have an effective, efficient Products Loss Control Program. A comprehensive Quality Control program starts with raw material evaluation and continues on through the entire manufacturing process and even includes packaging and shipping.

A statistical Quality Control program is desirable, but not absolutely necessary, as a function of the size of the insured (manufacturer), if the insured (manufacturer) has developed special quality requirements for critical parts.

An insured's Quality Control Department must have the following characteristics:

a) Independence - The Quality Control Department should report at a level equal to the production, engineering, and purchasing departments and should operate on a budget that is not part of another department's budget.

b) Stature - The Quality Control Department should have sufficient management stature, responsibility, and stability to establish and maintain an effective Quality Assurance system. The Quality Control manager should participate in top management meetings and decisions.

B. DEVELOPMENT OF TECHNICAL SERVICE CAPABILITIES

The desire to provide quality information to the underwriter has prompted the development, at least internally at INA, of a very strong products liability technical service capability.

What we have basically done is to develop engineering graduates into qualified Products Liability/Products Loss Control Consultants. It has been suggested that few manufacturers receive counsel on product safety from insurers because few insurers are able to retain engineers as qualified as the manufacturer to evaluate the safety of his product. This suggestion has been accepted and as far as INA is concerned we are responding.

Thus far we have hired 17 engineers from diversified backgrounds (electrical, mechanical, chemical, pipeline and natural gas, industrial, etc., engineering). These highly qualified individuals have been thoroughly trained via:

1) A Head Office Training Seminar

2) Correspondence courses

3) Development of an in-house Quality Control/Products Liability oriented manual for their use

4) Attendance at a Total Loss Control Seminar at the International Safety Academy

5) Field training by Head Office consultants

New, comprehensive forms have also been developed to assist them in their regionally oriented activities. In addition, detailed performance auditing techniques have been implemented in the Head Office in order to provide control and guidance to their development as Product Liability Consultants.

This program has just recently been expanded from 12 engineers to 17 and we may increase their number to 22 within the next two months.

C. PRODUCTS LOSS CONTROL PROGRAM IMPLEMENTATION

The Product Liability/Products Loss Control Consultant has been carefully selected and trained for a two-fold purpose.

1) The provision of Product Liability Survey service to both new and prospective insureds - Only a man who is well trained and experienced in the application of a superior products survey format will be able to achieve the primary goals of the survey, which are:

   a) The determination of all potential product liability exposures

   b) The evaluation of the insured's capability and desire to control exposures (in short, his Products Loss Control Program)

   c) The recommendation of adequate corrective measures when necessary.

   d) The motivation of the insured to comply with necessary recommendations.

   e) The accurate transmittal of pertinent survey information to the underwriter.

The Products Liability Consultant must be capable of achieving these goals in order to first, recognize a satisfactory risk; second, be able to improve what might otherwise be an undesirable risk; and third, aid the underwriter in his risk evaluation and rating procedure.

2) The provision of Product Loss Control consulting expertise to both insureds and non-insureds - Although, in today's social-legal-economic-political environment, any manufacturer/insurer could lose catastrophic amounts of money in products liability lawsuits arising out of the design, manufacture, and sale of products, the situation is not hopeless. Manufacturers can do much to control their products liability exposure through the implementation of good Products Loss Control Programs. Indeed, the existence of such programs will very likely mean the difference between a good and a poor risk as far as INA is concerned. Consequently, it is imperative for the Products Liability/Products Loss Control Consultant, during the course of a Product Liability Survey to evaluate not only the accident potential of the products manufactured and/or sold by the insured but also the insured's ability to effectively control this accident potential through good management techniques.

INA's Marketing Operations Division, Product Liability/Product Loss Control Consultants are trained not only to do both evaluations but also to assist the manufacturer (distributor, retailer, etc.) in improving his Products Loss Control Program to the point where the insured is capable of and has the desire to design, manufacturer, and sell reasonably safe products.

D. PRODUCTS LOSS CONTROL PROGRAM A NECESSITY

The success of a Products Loss Control Program within an industrial organization requires the presence of two factors. First, the influence of very strong Quality Control and Reliability programs must permeate all phases of product design and development. Second, and of paramount importance, the existence of a sincere, total commitment by company management to product safety is imperative. A commitment of this type will create the interdepartmental cooperation needed within a company to produce a completely effective Products Loss Control Program. The insurance industry needs and expects to receive this commitment from industry management in order to achieve its objective which is the profitable provision to industry of satisfactory products liability coverage, rates, and services.

We insist upon the implementation of an adequate Products Loss Control Program because from experience we know that a program of this type will provide reasonable control over the major sources of industrial products liability exposure which are:

III. MARKETING

Because liability for a product can be created by the manner in which the product is represented and marketed (even with reasonably safe products), sales brochures, instruction books, advertising, servicing agreements, etc., representing the product to the customer must be adequate, accurate and reasonable. Undesired warranties should not be given by the advertising and sales copy.

IV. RECORD KEEPING

Even with reasonably safe product designs and good manufacturing and quality controls, it is possible for batches or lots of defective products to slip through and reach the customer, creating serious products liability exposures. If this occurs the insured must have adequate records in order to be able to identify defective products and locate the customers who purchased them. Without adequate records, there is little the insured can do to control products liability exposures from defective products once they have reached the customer. Therefore, the insured's records must be comphrehensive (Design, Manufacturing, Quality Control, Sales and Service, etc.) as well as accurate if he is to create a favorable impression upon a jury if a products liability exposure does occur.

V. COMPLAINTS/INCIDENTS/ACCIDENTS

The insured's attitude toward consumer complaints, incidents, and accidents involving his products is extremely important. He must attempt to have the complaint/incidents/accidents reported by his traveling company personnel. They should then be analyzed and investigated for validity and potential safety hazards. The complaint and incident records must be evaluated regularly to determine the existence of trends that could possibly result in future products liability accident claims. And of most importance, if these activities are to be meaningful, the insured (manufacturer) must take action with respect to his findings.

VI. PRODUCT DESIGN

In the product design area the insured should

a) Meet or exceed all applicable safety standards.

b) Provide adequate warnings and instructions for safe use of the product.

c) Identify critical parts.

d) Completely analyze the product for accident hazards (Design Review).

I will now spend a little time discussing those Design Review related reliability techniques that the Products Loss Control Consultant has been trained to evaluate in order to determine the extent to which the manufacturer is actually committed to the development of safe, reliable products.

PREDICTION TECHNIQUES

Where appropriate the PLC Consultant will evaluate the scope of an insured's failure rate analysis (prediction) activities. Depending upon the type of product involved, the insured should be assigning estimated failure rates to each part contained in the product. They should be realistic and based upon an estimate of the failure rate of the part when under the stresses the product will be subject to when in use. The sources of failure rates utilized by a risk are also investigated. They can be obtained from a number of sources including: parts manufacturers, failure rates assigned by certain customers, historical information in the files of the designer of the part, etc. Major sources of failure rate data are available through the Library of Congress and U.S. Military procurement agencies. Electronic failure rate data are found in MIL-Handbook 217A. Farada data includes electronic, hydraulic, and mechanical failure rate data. Non-electronic data can be found in the USAF Reliability Handbook.

A measure of the risk's desire to accurately predict the overall failure rate of his product is the amount of effort he expends in the determination of unreliable segments of his product's design. In addition to reducing his potential products liability exposure from product failures the risk is advised that a thorough failure rate analysis will enable him to save money by accurately determining a satisfactory term and the estimated cost of providing the warranty.

TWO SYSTEMS - ANALYTIC APPROACHES:

FAULT TREE ANALYSIS AND FAILURE MODES AND EFFECTS ANALYSIS

1) FAULT TREE ANALYSIS

The Fault Tree was so named because the completed graphic delineation of a product or functional system looks like a (coniferous) tree. The undesired event is located at the top, or apex, and the various contributing events are the branches that extend laterally and down, from the top undesired event down to the least likely contributing factors. The Products Liability Consultant will evaluate the development of the fault trees in order to determine whether or not they are, in fact, comprehensive enough with respect to the product involved to ensure that all potential product liability exposure areas have been identified. A comprehensively done Fault Tree Analysis enables the risk to measure product safety and reliability because, theoretically, all potential events have been enumerated and every potential event has an associated probability of occurrence.

The Products Liability Consultant will determine whether or not the key to Fault Tree Analysis, the definition of the terminal event, or the event that is most undesired, has been satisfactorily accomplished. If this has been done properly, the terminal or undesired event can be designed away from so as to achieve product safety.

2) FAILURE MODES AND EFFECTS ANALYSIS

In this analysis each critical part in the product design is reviewed to determine what effect on the rest of the product design a failure of the critical part will have. The Products Liability Consultant will evaluate the FMEA to determine whether each potential mode of failure has been considered in the analysis.

A FMEA will analyze the effect each failure mode will have on the rest of the product. Based on this review, new parts can be specified, or if necessary, parts can be redesigned so failure will have a minimal effect on the product's liability (safety) exposure potential of the product(s).

We are particularly interested in the FMEA because it is particularly well-suited, in products liability work, to pin-pointing unexpected weaknesses in a product. They are normally found early in the development phase thus permitting the engineer to modify the design, change parts, processes, or materials in order to rectify the problem.

SUMMARY

The effects of the current products liability environment upon the Insurance Company of North America as reflected by the internal development of approaches to be utilized and skills to be employed in the solution of the problem have been expanded upon.

The highlights of the Products Loss Control Program have been outlined. This program is a tool that we demand an insured adequately employ as a condition of insurance, if necessary, because we are convinced of its efficacy in controlling products liability exposures. I might also add that it is a separate technical service that we also gladly sell to non-insureds.

Finally, I have placed some emphasis on the role that reliability technology plays in the overall Products Loss Control Program picture. It is a sophisticated technology and, as such, occupies a prominent place in the Product Liability Consultant's scope of analysis.

To be perfectly frank it is obvious that we must place a certain amount of reliance on the reliability engineers willingness to assist us in our evaluation of the adequacy of his Products Liability oriented reliability analyses. To this point in time we have not been disappointed.

REFERENCES

1. Libertiny, G. Z. and D'Souza, A.F., Product Safety: Is the Engineer Morally Bound?, ASME Winter Annual Meeting, December 1970.

2. Shankula, R.E., Products Loss Control Guide, Insurance Company of North America, March 1970.

3. Marhefka, R.E., Product Liability: An Insurer Responds to the Challenge, Transactions of the Ninth Annual Conference and Technical Exhibit on Quality Control and Reliability, Stony Brook, New York, April 1971.

4. Final Report of the National Commission on Product Safety, June 1970, Part III, Chapter 6.

5. Jacobs, R.M., Minimizing Hazards in Design, Proceedings of the Product Liability Prevention Conference, Newark, N.J., August 1971.

6. Selman, J.H., and Selman, V., Two Approaches for Liability Risk Evaluation, Proceedings of the Product Liability Prevention Conference, Newark, N.J., August 1972.

THE FORMULA FOR SURVIVAL -- OPTIMUM QUALITY AT OPTIMUM COST

Albert Boerckel, Jr., Chief Inspector, Quality Control General Offices
Caterpillar Tractor Co., Peoria, Illinois

INDEX SERIAL NUMBER – 1095

The U.S. industrial complex is today confronted with a variety of challenges. We might quickly pass this off by saying, "What else is new?" I'm convinced that these are challenges of deeper concern and a result of a changing consumer attitude, steady ever-changing social values, competition, and higher costs.

The strength of the dollar and the quality and performance of U.S.A. products may not be of sufficient stature to maintain a preferred position in the world marketplace of the future.

European, Asian, and South American industrial growth is accelerating as they strive for worldwide commercial integrity, product quality, and self-respect with an aggressive drive for national dignity and recognition in global markets. In order to meet and surpass this relentless challenge from abroad, U.S. industry will have to do things differently. The business-as-usual approach will not be sufficient, and the current practices, policy, and philosophy must be reviewed.

We witness products being produced in new or modernized industrial areas in countries where a lower standard of living prevails -- thus providing the opportunities to manufacture, ship halfway around the world, and market at very competitive prices.

As these countries continue to progress through dedication, application, and ingenuity, they have discovered some most effective methods of producing large volume, reliable products. We also witness their emphasis on the growth of GNP with the social and ecological problems being relegated to a lower priority.

Foreign governments have also participated by aiding their country's business in several ways -- first, by elevating the priority for quality on a national scale and, second, by encouraging industry to use quality as a competitive edge. Japan, for example, certainly had a past reputation for substandard quality products. This is not the case in today's marketplace. In fact, their product now carries with it a most respectable quality reputation.

We find ourselves in an almost impossible position of being unable to cope financially with a completely new and modern industry. I suggest we seek alternatives and adjust our approach to the needs required to bring about a successful, aggressive response to the challenge before us.

We must be willing to accept the idea that we do not always have the one-and-only or best idea in the world of today. It is therefore essential that we become good listeners. Deadlines for product shipment are currently becoming less important. Quality of product acceptable to the consumer will need to take a predominant role in the decision whether or not to ship.

Uppermost in our minds should be quality and cost.

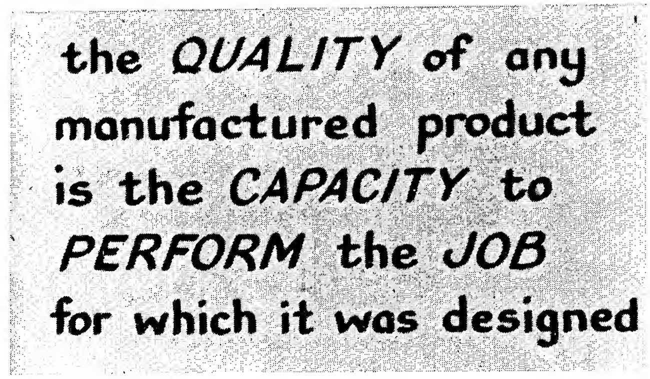


Figure 1

Each of us has a concept of what "quality" means, and it is most important that I state a definition which I feel is most realistic.

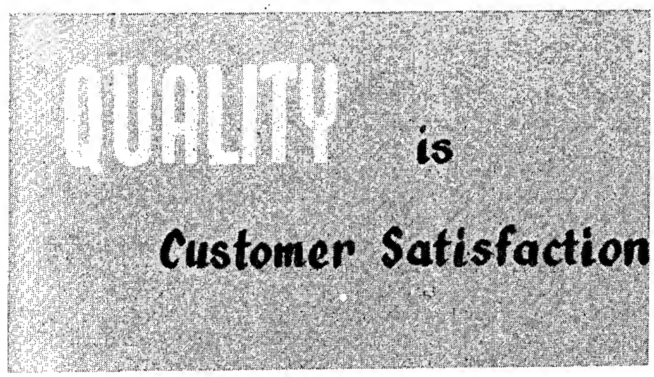


Figure 2

The quality of any manufactured product is interpreted as "the capacity to perform the job for which it was designed." If a customer has a complaint about a product, he will comment that it lacks quality; therefore, "quality is customer satisfaction."

It is quite easy to relegate quality problems to a lower priority; therefore, it is essential to review where quality lies in our priorities. Manufacturing today must assure themselves that quality rests on a par with profit, for without one the other won't last for long. We are in business to make money, and obviously this is the main objective of a company or corporation. This kind of thinking must emanate from the top of the organization, whether it be a small company or a corporate level of a multinational company.

As the chief executive officer sets down the short- and long-range objectives and budget guidelines, the quality control organization must also establish the needs and basic operating organizational structure required to fulfill these needs. A simple basic organizational chart is illustrated here.

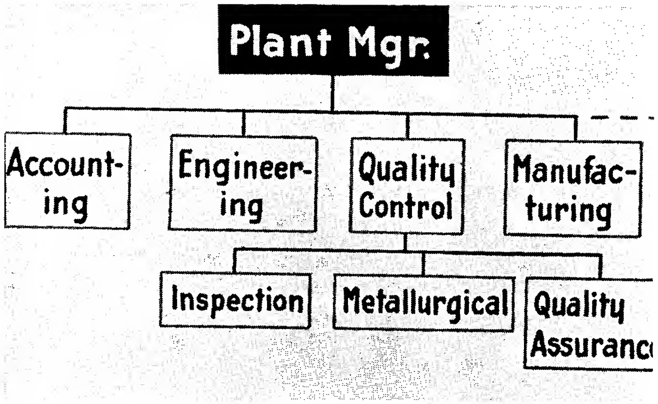


Figure 3

Although this is a very simple organization, it does set up the framework from which an aggressive independent quality program can emanate.

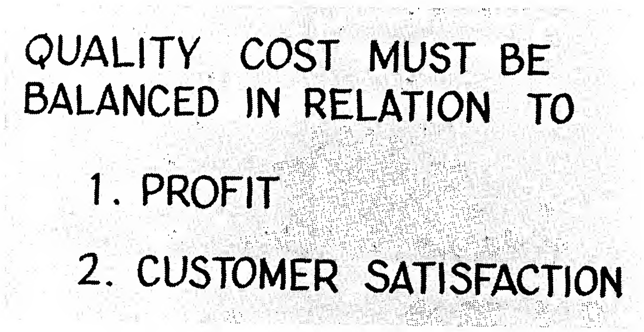


Figure 4

The quality dollar being spent wisely today will not only improve the present profit picture but will result in substantially fewer field customer complaints and warranty costs tomorrow.

To effectively budget the quality dollar, quality costs must be defined and quantified.

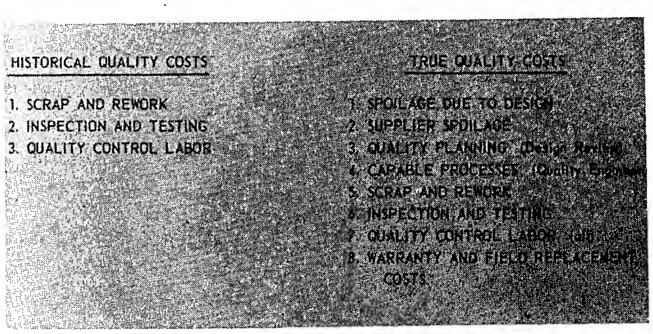


Figure 5

True quality costs are sometimes difficult to obtain because generally company accounting systems are designed by accountants for accountants, and costs are so coarsely grouped that they are difficult to use as a management tool. To develop true quality costs, we must begin with <u>four</u> basic categories:

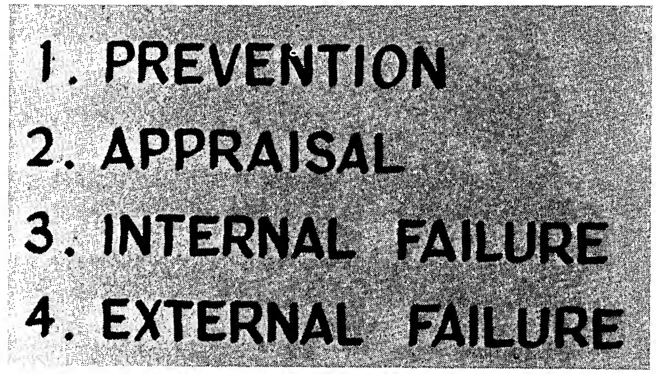


Figure 6

Into these categories we will fit the eight quality costs previously listed.

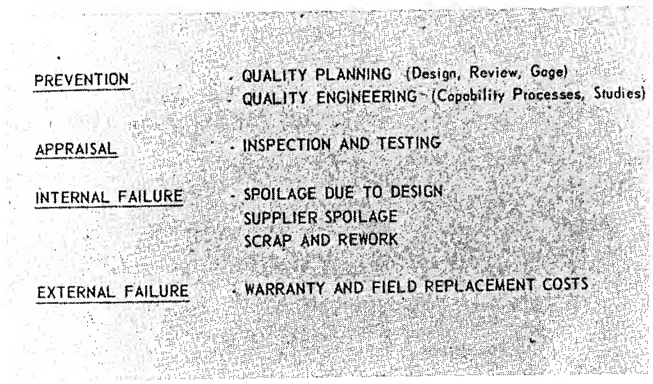


Figure 7

The total cost of the four categories must be measured against sales and expressed as a percent.

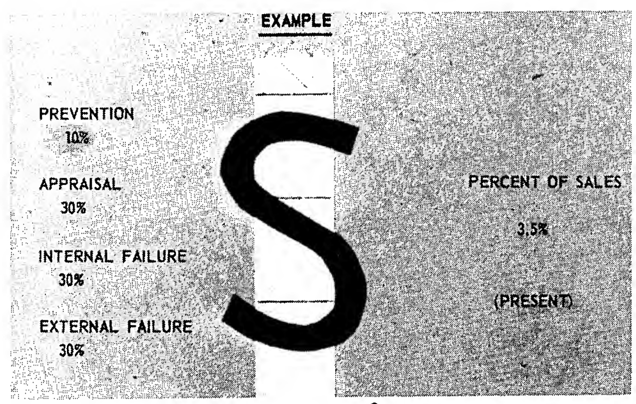


Figure 8

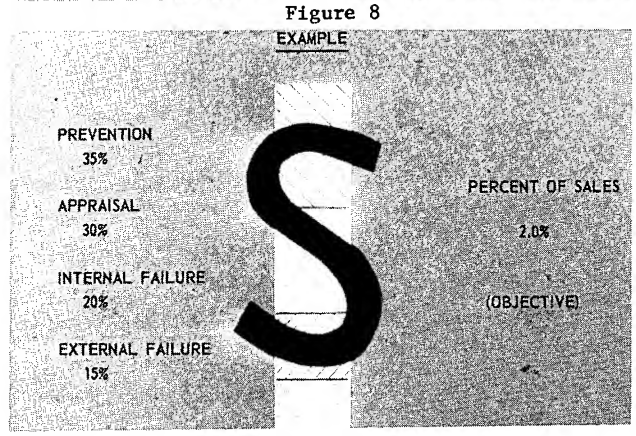


Figure 9

In reviewing the two charts showing the allocation of the quality dollar, two things are apparent. When the accent is placed on prevention:

1. internal and external failures decrease, and
2. as a result, total quality costs are lower.

There are other benefits achieved that reflect in the corporate picture.

1. Customer acceptance is improved.
2. Manufacturing efficiency is higher due to fewer lost hours (scrap and rework hours).
3. Material flow is improved (less stock sorting and reordering).

Let's review the differences between the two quality systems. The system lacking prevention and accountability is geared to react to events that have occurred.

An example is the introduction of a new product.

1. DESIGN CREATED
2. MACHINE AND TOOLS PROCURED
3. ROUGH MATERIAL PROCURED
4. PARTS MANUFACTURED ← START QUALITY CONTROL ACTIVITY
5. ASSEMBLY AND TEST
6. SHIP

Figure 10

With quality control activity starting at the manufacturing stage, all that is accomplished is the sorting of nonconforming product caused by system errors.

A. Expensive materials
B. Incapabile processes (heat treat and machining)
C. Excessive customer complaints/high warranty.

The prevention approach starts at the design phase with drawing review for functional tolerancing and drawing clarity.

1. DESIGN CREATED ← START QUALITY CONTROL ACTIVITY
2. MACHINE AND TOOLS PROCURED
3. ROUGH MATERIAL PROCURED
4. PARTS MANUFACTURED
5. ASSEMBLY AND TEST
6. TEST

Figure 11

Emphasis has been placed on quality starting at the top with high priority and related cost. To implement these basics, there are four principles to be accepted.

1. QUALITY MUST BE PLACED AS ONE OF THE PRIME ONGOING CORPORATE OBJECTIVES.
2. THE QUALITY EFFORT MUST BE GEARED TO OPERATIONAL COST REDUCTION.
3. THE ROLE OF QUALITY MUST BE SO IMMERSED IN CORPORATE PLANNING THAT IT IS CONSIDERED IN EVERY FACET OF ORGANIZATION AND OPERATION.
4. ALL ACTIVITIES NEED TO BE EITHER IMPROVED OR ELIMINATED.

Figure 12

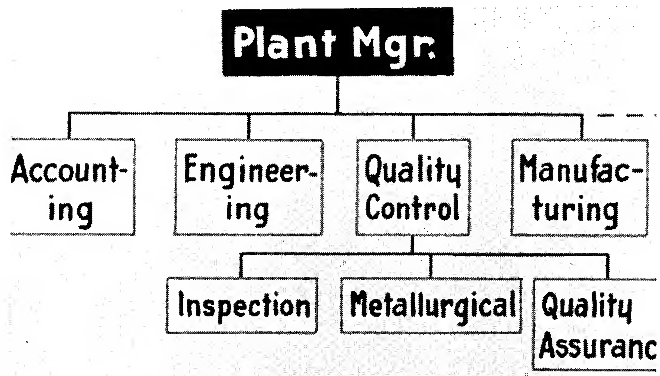


Figure 13

In the simple organizational chart shown earlier, Metallurgy is a separate division of the Quality Control Department and a part of the quality team effort responsible for results. Metallurgy has been a very important partner in the quality and cost discussion as it seeks its identity through careful quality planning at the proper stage of development in decisions on type of material, control of that material for cleanliness requirements, machinability, type of process, fatigue properties, etc. These must be finalized to assure a capable process at optimum cost for customer satisfaction.

1. QUALITY MUST BE PLACED AS ONE OF THE PRIME ONGOING CORPORATE OBJECTIVES.
2. THE QUALITY EFFORT MUST BE GEARED TO OPERATIONAL COST REDUCTION.
3. THE ROLE OF QUALITY MUST BE SO IMMERSED IN CORPORATE PLANNING THAT IT IS CONSIDERED IN EVERY FACET OF ORGANIZATION AND OPERATION.
4. ALL ACTIVITIES NEED TO BE EITHER IMPROVED OR ELIMINATED.

Figure 14

It is in applying the second of these principles to the field of metallurgy that the goal becomes the optimum combination of material and processing to fill the intent of the design of the component. It requires no special talent to overspecify material or processing

(or both) to completely avoid any risk of product failure to fulfill the design. This approach serves the metallurgist well but projects very high risks to long-term customer acceptance and corporate profitability. Likewise, the meat-axe approach to cost reduction in materials and/or processing requires little technical competence and no great amount of intellectual effort.

This extreme spells disaster for customer acceptance and corporate profitability.

Achieving the goal of the best combination (or combinations) of material and processing is the risk-free route to customer acceptance and corporate profit improvement. It involves a great deal more effort, anxiety, ingenuity, and so on, but it also offers the ultimate to personal challenge and potential for reward.

Definition of actual engineering requirement is a critical starting point. More often than not, design is based on precedent--not calculation. This complicates the definition of actual engineering requirement and tends to perpetuate the status quo. On the other extreme, it provides strong temptation to "reinvent the wheel" where the payout may be counterproductive.

Solid data, in-depth analysis, and clear communication with the designer are the keys to success in this aspect of materials selection.

Selection of best material and processing is one step in achieving a competitive metallurgical position. Maintaining this position involves a system and related activities of another sort. An important factor in the selection decision is the process capability of the material/process combination. This is similar in all respects to the same factor in machining processes, and in addition it must be measured, monitored, and reacted to throughout the production span of the application with intensity modified as the status changes.

It involves measurement of variables in the process and variation in the product as well. Such factors as temperature cycles, time cycles, furnace atmosphere composition in a carburizing process, etc., must be measured and monitored, preferably by automatic means since human performance is rapidly becoming the weakest link in the process control chain.

Identification of the significant characteristics of the product that directly correlate with the engineering requirements is a basic part of process capability analysis, and then the variations inherent in the material/process combination chosen have to be established. This requires sampling on a rigid statistical basis measurement of the significant variations and computation of the total process capability. Anything less than this is only random sampling--which has little value as an indicator of real process capability.

It is appropriate at this point to discuss an example in which material and processing (heat treat and machining) contributed to the existence of an incapable process producing high scrap and rework.

In this example the product material was specified as SAE 1048, then processed through a series of rough and finish machining operations with a final tolerance of ± .00025 outside diameter specified. This tolerance was required prior to the subsequent heat treat operation consisting of a heat treat cycle, chemical treatment, and final quenching operation. The distortion, as a result of heat treat processing, must be controlled to a ± .00025. The material and process was initially selected to provide a component with torsional and high bending fatigue strength at an economical cost.

The result of this decision was a very costly operation consisting of high scrap and rework accompanied by the inherent high processing costs, particularly in heat treatment.

Because of the unsatisfactory results and high cost, a new process was developed which subsequently eliminated this high scrap and rework condition. The new process consisted of an intermediate heat treatment of a different type, which resulted in a change in tolerance from ± .00025 to a ± .0005. In addition the process was designed to produce at a faster rate than could be accomplished by the initial design. Also, the new process created the ability to change to a new and less costly material. The fruits of these efforts have resulted in a capable process, at less cost and with improved quality, in addition to being an improved product for the customer.

The capability analysis is as follows:

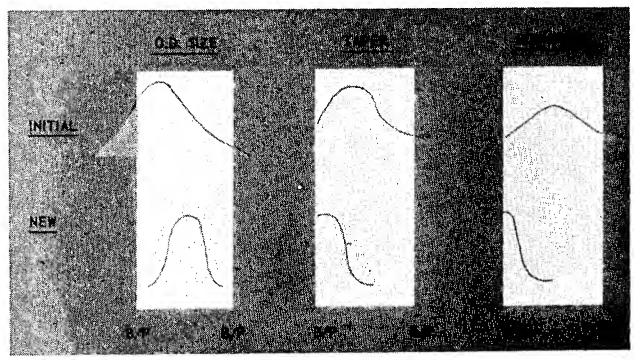


Figure 15

1. QUALITY MUST BE PLACED AS ONE OF THE PRIME ONGOING CORPORATE OBJECTIVES.
2. THE QUALITY EFFORT MUST BE GEARED TO OPERATIONAL COST REDUCTION.
3. THE ROLE OF QUALITY MUST BE SO IMMERSED IN CORPORATE PLANNING THAT IT IS CONSIDERED IN EVERY FACET OF ORGANIZATION AND OPERATION.
4. ALL ACTIVITIES NEED TO BE EITHER IMPROVED OR ELIMINATED.

Figure 16

In applying the third principle, it is necessary to consider the role of the remaining two divisions within Quality Control--Quality Assurance and Inspection.

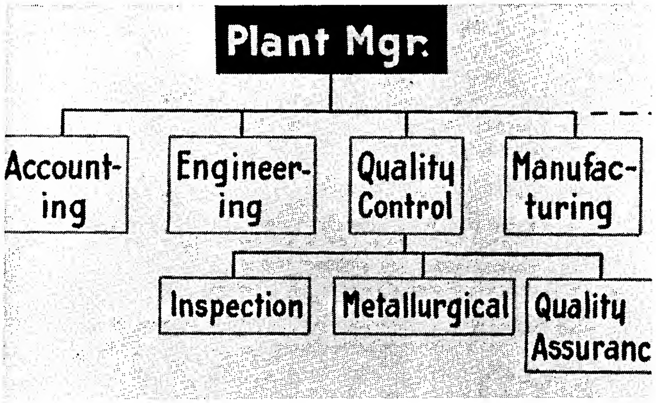


Figure 17

Inspection is the on-the-spot monitor of processed materials and product, an extremely important activity. However, since the prime objective is to emphasize the planning and capability portions, and as stated earlier under organization, I'll refrain from going into detail here also.

In introducing the subject of total process capability, it is necessary to discuss the Quality Assurance role as a vehicle/catalyst to provide the creative environment necessary for this, the engineering and planning arm within Quality Control.

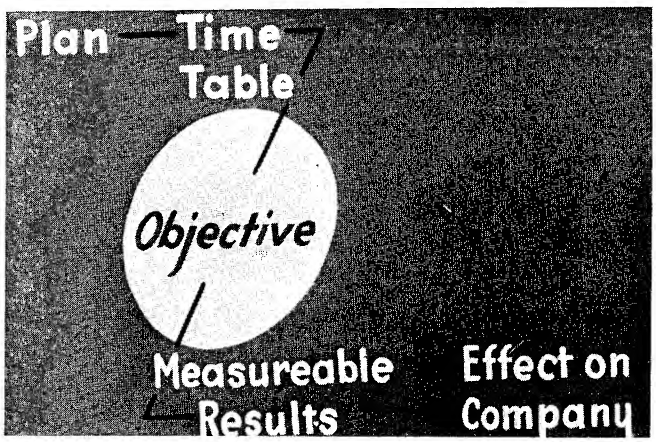


Figure 18

It is through the activities of this group that guidelines and procedures for quality planning and activity begin at the earliest stage of the product. A typical modern version of this organization can be:

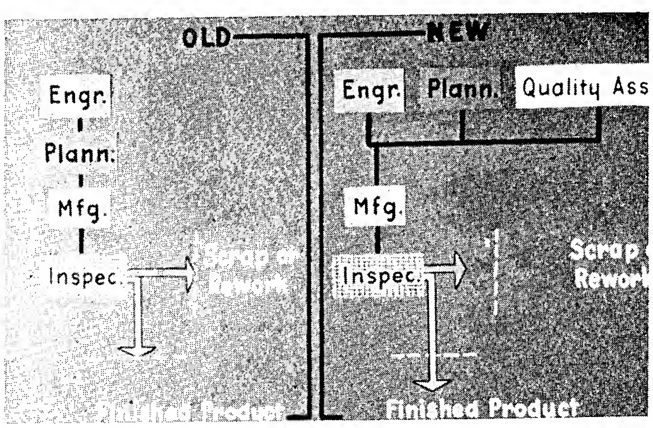


Figure 19

Activity must start in the design phase by defining quality parameters for proposed product and by agreeing upon how these are to be evaluated.

This is a broad team effort consisting of Quality Assurance, Metallurgy, Engineering, Manufacturing, Planning and Tooling, Purchasing, etc. Their role is to require a review of each dimension as well as metallurgical needs and specifications testing their capability to produce to each tolerance, including supplier capabilities, where required, to produce rough and finished materials. This may require supplier input in cases of complicated design of material.

Once the team members concur, machine tools and processes can be finalized for producing to a predetermined dispersion.

Process control is the combination of many variables: type of material, cutting oils and tool grade cycle time, hardness of material, etc.

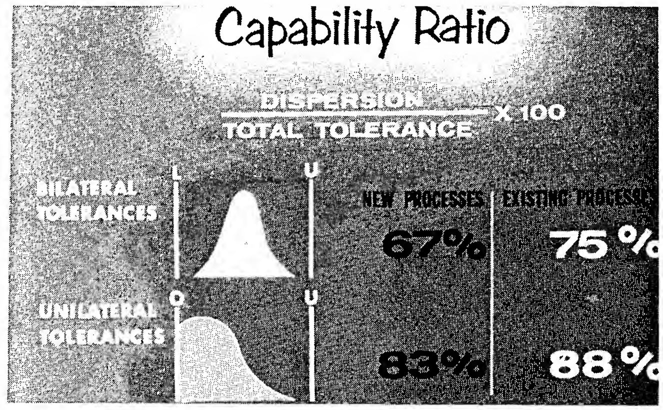
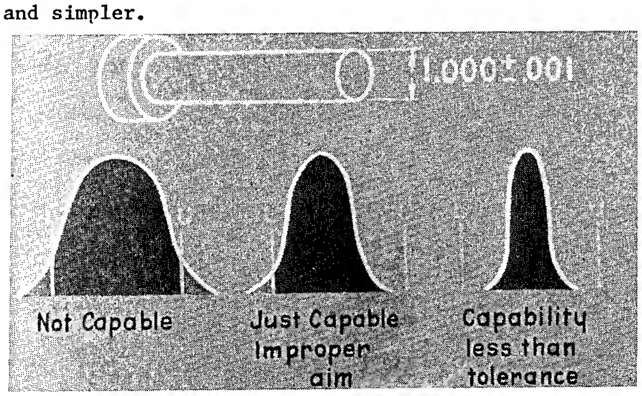


Figure 20

Important at this point is the trial run or runoff specification for the process. This must be defined to enable bidding by machine tool, wash tank, heat treat furnace, and special process suppliers -- the eventual goal being a successful runoff of each process at the supplier with purchase of the equipment only after compliance to tolerance capability requirements.

This approach is but common sense. Imagine purchasing a machine or process without evidence of its actual ability to produce to a capability dispersion centered within the specified tolerance! Once in your plant, it is late to start debugging an incapable process. The cost of this is tremendous to both your company and to the supplier.

The statistical means have been available for years to compile results. Computer programs to evaluate the dispersion produced have made it even faster and simpler.



We can learn about material, grinding of tools, coolants, machine tools, etc. Information from all sources can be compiled in data banks for rapid retrieval to be used by engineering standards, designers, planning processors, time standards, etc. This serves to shorten the time on the next review. Imagine having a bank of actual data from your worldwide organization available at your finger tips. The ultimate goal is optimum quality at optimum cost.

The monitoring of the process for control as utilized in metallurgical processing is of joint importance. It is this combination of all the elements inherent in a process, which has been carefully planned, through which you provide capability and meet the engineering requirements and assure reliability.

Concentration thus far has been on new product and new processes. This brings us to current product and processes. What about them? Are they capable? Are operating costs high?

1. QUALITY MUST BE PLACED AS ONE OF THE PRIME ONGOING CORPORATE OBJECTIVES.
2. THE QUALITY EFFORT MUST BE GEARED TO OPERATIONAL COST REDUCTION.
3. THE ROLE OF QUALITY MUST BE SO IMMERSED IN CORPORATE PLANNING THAT IT IS CONSIDERED IN EVERY FACET OF ORGANIZATION AND OPERATION.
4. ALL ACTIVITIES NEED TO BE EITHER IMPROVED OR ELIMINATED.

Figure 22

The fourth principle considers a review of all activities which need to be either improved or eliminated.

Competition, economic conditions, and the world market demand cost reduction activity. Where is a better place to start than by reviewing the current processes for capability? This can begin by reviewing the scrap and rework reporting systems, followed by a systematic approach utilizing the "Parento" curve concept of placing high loss items in dollar order.

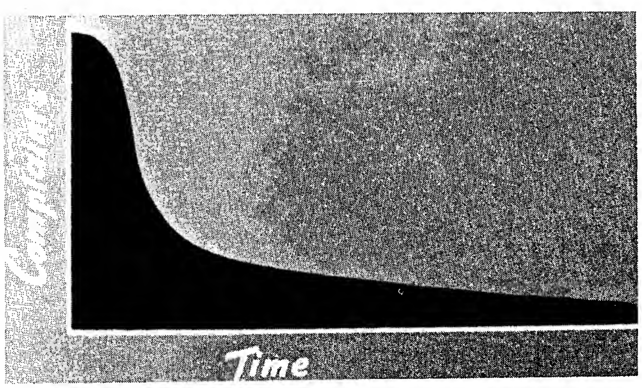


Figure 23

Attack the highest loss items by process capability study after the process has been reviewed for correctness to plan. This will reveal true nonconformance causes and provide capability information for action by Metallurgy, Manufacturing, Quality Control, Engineering, etc.--"the team." This affords an opportunity to review design specifications and to bring about corrections based on history and fact. Consideration should be given to material selection. This could become a major portion of profit improvement by correcting past mistakes.

The providing of new, less costly material combined with improved processes resulting in reduced scrap and rework will, from experience, give a fair return of from $14 to $10 for $1 expended in this type of review activity. In addition the necessary discipline can be devised to assure a continued accurate process monitoring.

In moving on to complete the cycle of design and process control, it is not sufficient to merely assemble the product and ship it; it is essential to test the completed product for reliability. A test program must be developed, scheduled to begin with assembled product, prior to normal production activity. This test should be designed and followed by Quality Control in conjunction with Research and Engineering. It should be a separate test from the normal Research and Engineering test program. The test should be planned and conducted as though you were the customer. This is a test for design capability and should be conducted well in advance of scheduled production. This allows the time necessary for corrections. Deficiencies must be recorded carefully and documented to guide corrective action. Retesting is the only assurance that corrections are effective. Testing duration should be of extreme importance. This is an accelerated type test to reveal early design failures and thus assure reliability to the design.

As production starts, further testing is required. Testing would consist of a random selected unit from first production. As this is also an accelerated type test, its main purpose, however, is to prove out the manufacturing, assembly component testing, and tooling techniques, including proper adjustment, leaks, etc. No shipment can be made to customers until this test has satisfactorily passed the requirements and fixes, if required, are completed. Usually production build consists of a very few units with a stop order until test results are complete. A sound fast reaction type field or customer feedback system is essential for action on field problems developing after certain time spans under varying circumstances. A reliable method of doing this can be through the service arm of the business and by a controlled number of units being introduced to the field through customers that will obtain maximum use in a short period of time. This provides day-to-day surveillance of the product with immediate feedback. Although merely a one-shot type of operation that cannot be accomplished throughout the life of every unit produced, the system must in addition be so designed as to provide for regular feedback through customer, dealer, and factory communication.

Warranty is essential to the business and requires a well-planned realistic policy. Warranty can be a measure of customer satisfaction and a measure of product reliability.

In summary--What is the effect? What will change? What can we achieve with optimum quality at optimum cost?

A quality attitude will become engrained throughout the company. Quality will improve at a lower cost to help assure survival in the world marketplace.

The improved profit margin will be most beneficial to the total enterprise.

In addition, I suggest the following:

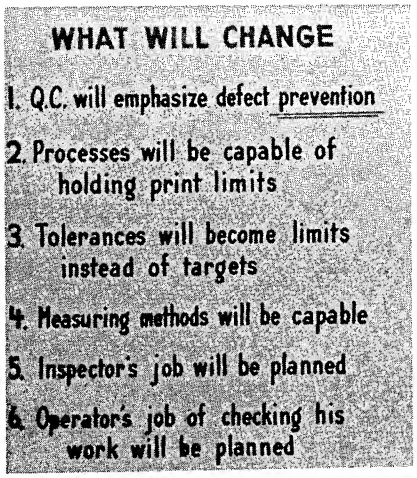


Figure 24

This is not some magic wand to wave, but a lot of very hard work to meet the challenge before us and to assure reliability of product for worldwide customer satisfaction.

A STATISTICAL ANALYSIS OF PRODUCT RELIABILITY DUE TO RANDOM VIBRATION

INDEX SERIAL NUMBER — 1096

Arthur Sorensen, Jr.
Associate Professor of Mechanics
College of Engineering and Applied Science
University of Wisconsin-Milwaukee
Milwaukee, Wisconsin

Summary

The durability of equipment which operates in a random service environment often depends upon the fatigue life of mechanical components. A systematic investigation of the fatigue damage induced by the random vibration is necessary to assess the product reliability under these conditions. The purpose of this study is to examine the various links in the statistical chain between the random vibration and the product reliability in a deliberate, logical manner. This presentation considers the role of each link in this chain, and it pin-points the gaps that exist in the present state-of-the-art. This leads to a more precise statistical analysis of random fatigue, which is indispensable to a more satisfactory understanding of this elusive topic.

Introduction

It is possible to examine material behavior at various levels of observation - from the discrete, local aspect of material science to the continuous global perspective of engineering mechanics. The former is conducive to a fundamental study of material behavior, while the latter facilitates a logical approach to mechanical design. From a practical viewpoint, the progressive fracture mechanism consists of a preliminary deterioration of the material which initiates a visible crack whose subsequent growth may result in eventual disintegration of the adjacent structure. It is expedient to classify the mode of failure on the basis of strength (in the case of ductile flow or brittle fracture), which depends upon the current stress level, and life (in the case of creep rupture or fatigue failure), which depends upon the previous stress history. In the former "imminent" mode of failure, immediate fracture is supposed to occur at a critical stress level, while in the latter "cumulative" mode of failure, crack incubation is presumed to ensue from damage accumulation. In contrast to time-dependent creep under constant load conditions, cycle-dependent fatigue is contingent upon a repetitious load variation.

A distinct fatigue fracture mechanism is operative in the high level (low cycle) range above the yield point and the low level (high cycle) range below the yield point, respectively. In the former situation, the region of plastic deformation is quite extensive, and the fracture pattern is ductile in nature, whereas the plastic zone is confined to the immediate vicinity of the crack, and the fracture pattern is brittle in appearance in the latter situation. The activation of both of these fracture mechanisms may be induced by a load variation, over a wide amplitude range (or acute stress concentrations which cause plastic strain variations) with the possibility of a mutual interaction between them. In order to limit the scope of this presentation, the following discussion refers only to the ordinary type of fatigue failure, which is due to a stress variation in the intermediate range between the endurance limit (below which the fatigue life is indefinite) and the yield point of the material.

Material Behavior

The fatigue failure mechanism results in eventual formation of a visible crack at a critical location in a structural member due to repetitive application of a variable load.

Consider a simple harmonic stress variation S(t) as follows:

$$S(t) = a \sin \Omega t \tag{1}$$

where a is the amplitude, Ω is the frequency, and t is the time.

It is possible to describe the fatigue life L by a simple power law of failure[1]:

$$Na^{\beta} = \lambda \text{ with } \Omega L = 2\pi N \tag{2}$$

where N is the number of cycles to failure, β is the fatigue exponent, and λ is another material parameter. This behavior, which is depicted in Fig. 1, is independent of the excitation frequency.

The above relation provides a concise description of the general pattern of material behavior, which contains many exceptions that are difficult to incorporate into a simple theory of fatigue failure. Nevertheless, it enjoys a substantial degree of empirical validity under suitable conditions in the absence of a corrosive atmosphere or temperature elevation, which introduce time-dependent effects. The empirical constants (which are supposed to be independent of the stress variation in the intermediate range between the yield point and the endurance limit) allow for the influence of various factors, such as material composition, heat treatment, ambient temperature, and the like, which determine the specific nature of the failure characteristic.

Damage Accumulation

The basic fatigue life relation indicated by eq. (2) refers to a simple harmonic stress variation as stipulated by eq. (1), with a stress amplitude that remains constant prior to failure. Of course, this situation is rather artificial, since the load intensity may change quite often during the life of the member. Moreover, the load variation is not always harmonic or even periodic in nature, and the excitation may exhibit a very complicated waveform on some occasions. It is therefore necessary to devise a more general strategy to deal with sporadic changes in the load intensity as well as complicated waveforms due to irregular load variations.

The eventual formation of a fatigue crack is the visible outcome of the physical damage induced by a load variation, whose cumulative effect is responsible for this terminal result. This progressive damage accumulation evidently occurs in some continuous manner, and it apparently depends upon the cycle ratio C defined below:

$$C = \frac{n}{N} \tag{3}$$

where n is the actual number of stress cycles at a given stress level, and N is the number of cycles to failure under the same conditions. It is natural to suppose that the total damage D is related to the cycle ratio C_i at each stress level and to surmise that this might take the form of a simple linear combination as follows:[2]

$$D = \sum_{i=1}^{j} C_i = \sum_{i=1}^{j} \frac{n_i}{N_i} = 1 \text{ when } t = L \tag{4}$$

with failure incident at a critical value of unity. This linear damage hypothesis is just about "as good as any other theory" now available, and it is doubtful if the modest gain in accuracy provided by a more elaborate theory is sufficient to justify the serious extent to which it complicates the analysis. Despite its limitations, the above damage hypothesis is used extensively by the aircraft industry, and it is acceptable to the Federal Aviation Administration in practical design calculations.[3]

The damage hypothesis indicated by eq. (4) is the key to a more general theory of damage accumulation for the sort of irregular waveform depicted in Fig. 2. In this event, it is natural to replace the previous summation by a corresponding integral:

$$D(t) = \int_0^t R(u)\,du = 1 \text{ when } t = L \tag{5}$$

where D is the cumulative damage at the present time t, and R is the damage rate at any time u in the past. The form of the damage rate (which depends upon the stress variation) is not arbitrary, of course, since it must be compatible with the fatigue life relation given by eq. (2) when the stress variation conforms to eq. (1). This provides a basic guideline which may be used to determine the specific form of the damage rate in terms of the stress variation under more general conditions.

The fatigue damage induced by a harmonic stress variation is independent of the excitation frequency over a wide frequency range. It is possible to extend this basic concept of cycle-dependent behavior* to irregular waveforms if the damage rate is given by the time derivative of a suitable function F of the stress variation[4]:

$$R = \left|\frac{dF}{dt}\right| \text{ with } F = \frac{|S|^\beta}{4\lambda} \tag{6}$$

This yields the following expression for the damage rate:

$$R = \frac{1}{4\lambda}\left|\frac{d}{dt}|S|^\beta\right| = \frac{\beta}{4\lambda}|S|^{\beta-1}|\dot{S}| \tag{7}$$

where the superscript dot indicates a time derivative. A substitution of eq. (7) into eq. (5) then yields the fatigue life given by eq. (2) when the stress variation is given by eq. (1).

* Note: The prospect of some sort of alternative frequency domain analysis is intriguing, but it leads to serious (if not insurmountable) obstacles which impede progress in this direction.

The fatigue damage associated with irregular stress waveforms and peculiar stress conditions involves some rather subtle considerations which demand fundamental investigation. It is possible to avoid the more speculative aspects of this complicated subject if we confine our attention to a stress variation about a zero stress level. In this event, the consecutive peaks S' and dips S" in the absolute profile of the stress variation about the zero stress level depicted in Fig. 3 have a decisive influence on the cumulative damage irrespective of the intermediate waveform between these "critical" points as indicated below[5]:

$$D = \frac{1}{2\lambda}\sum_{i=1}^{j}\left(S_i'^{\beta} - S_i''^{\beta}\right) \tag{8}$$

This expression provides a logical basis for a statistical analysis of random fatigue and product reliability.

Product Reliability

The fatigue life associated with irregular stress waveforms may be calculated from eq. (8) in a deterministic situation. Of course, the stress may exhibit a random variation, in which event we must resort to a statistical analysis instead of a definite evaluation of the fatigue life. Under these circumstances, the basic objective is to deduce the stochastic behavior of the latter from a suitable description of the former in terms of certain probability distributions or statistical moments.

It is possible to relate the product reliability to the probability distribution of the maxima and minima of the irregular stress waveform depicted in Fig. 2. Let s(n,a,t) da denote the probability of n maxima or minima in the stress range $a \leq S \leq a + da$ over the time interval t as indicated in Fig. 4. This discrete probability density function consists of an impulse at each positive integer n with a variable intensity s that is a continuous function of the stress level a and the time interval t. The positive or negative damage increment d associated with each impulse is given by:

$$d = \frac{1}{2N} = \frac{a^\beta}{2\lambda} \text{ when } S = a \tag{9}$$

The probability density q(D,a,t) of the cumulative damage D is accordingly given by:

$$q(D,a,t) = s(n,a,t) \text{ with } D = nd = \frac{na^\beta}{2\lambda} \tag{10}$$

This corresponds to the density spectrum illustrated in Fig. 5, with an impulse at integral multiples n of the damage increment d. The intensity of each damage impulse is the same as the corresponding impulse in Fig. 4, while the location on the damage scale depends upon the stress level a as indicated by Fig. 5.

We may now consider the probability that the cumulative damage is less than a given value D and partition the stress level into suitable intervals[6]:

$$n = k \text{ when } \frac{ka^\beta}{2\lambda} < D < \frac{(k+1)a^\beta}{2\lambda} \tag{11}$$

The corresponding stress intervals are then given by:

$$\left(\frac{2\lambda D}{k+1}\right)^{\frac{1}{\beta}} < a < \left(\frac{2\lambda D}{k}\right)^{\frac{1}{\beta}} \text{ with } n = 0,1,2,... \qquad (12)$$

The total probability Q associated with the various stress intervals is then given by:

$$Q = \sum_{n=0}^{\infty} \int_0^{\left(\frac{2\lambda D}{n}\right)^{\frac{1}{\beta}}} s(n,a,t)da \text{ when } D < 1 \qquad (13)$$

We may account for the positive and negative damage increments as well as the positive and negative stress intervals as follows:

$$Q(D,t) = \sum_{n=0}^{\infty} \int_0^{\left(\frac{2\lambda D}{n}\right)^{\frac{1}{\beta}}} [s'(n,a,t)-s''(n,a,t)]da$$

$$+ \sum_{n=0}^{\infty} \int_{-\left(\frac{2\lambda D}{n}\right)^{\frac{1}{\beta}}}^{0} [s''(n,a,t)-s'(n,a,t)da] \qquad (14)$$

where s' and s" refer to the maxima and minima, respectively. A simple manipulation then yields the probability distribution function Q(D,t) of the cumulative damage D at time t as follows:

$$Q(D,t) = \sum_{n=0}^{\infty} \int_0^{\left(\frac{2\lambda D}{n}\right)^{\frac{1}{\beta}}} F(n,a,t)da \qquad (15)$$

where the function F is given by:

$$F(n,a,t) = s'(n,a,t)-s'(n,-a,t)-s''(n,a,t)+s''(n,-a,t) \qquad (16)$$

The probability $\rho(t)$ of survival to time t is accordingly:

$$\rho(t) = Q(1,t) = \sum_{n=0}^{\infty} \int_0^{\left(\frac{2\lambda}{n}\right)^{\frac{1}{\beta}}} F(n,a,t)da \qquad (17)$$

The probability P that the fatigue life is less than or equal to L is given by:

$$P(L) = 1 - Q(1,L) = 1 - \rho(L) \qquad (18)$$

The continuous probability density function p of the fatigue life L is given by the derivative of the distribution function P as follows:

$$p(L) = P'(L) = -\rho'(L) =$$

$$- \sum_{n=0}^{\infty} \int_0^{\left(\frac{2\lambda}{n}\right)^{\frac{1}{\beta}}} \frac{\partial}{\partial L} F(n,a,L)da \qquad (19)$$

The statistical moments μ_n of the fatigue life are then given by:

$$\mu_n = \int_0^{\infty} L^n p(L)dL =$$

$$- \sum_{m=0}^{\infty} \int_0^{\left(\frac{2\lambda}{m}\right)^{\frac{1}{\beta}}} \int_0^{\infty} L^n \frac{\partial}{\partial L} F(m,a,L)dL\, da \qquad (20)$$

Integration by parts then yields:

$$\mu_n = n \sum_{m=0}^{\infty} \int_0^{\left(\frac{2\lambda}{m}\right)^{\frac{1}{\beta}}} \int_0^{\infty} L^{n-1} F(m,a,L)dL\, da \qquad (21)$$

The product reliability is related to the probability distribution of the fatigue life and the cumulative damage by eq. (18), as well as the statistical behavior of the maxima and minima of the stress waveform through eq. (17) and eq. (16). Unfortunately, the latter is unknown in many situations, and it is not easy to obtain this information from a statistical description of the random excitation. Although the cumulative damage and the damage rate are related by the stochastic integral of eq. (5), the statistical connection between these random variables is difficult to establish. This is the weakest link in the statistical chain between the product reliability and the random vibration.

Statistical Analysis

The computational accessibility of the damage rate from the stress variation indicated by eq. (7) is a very important practical consideration in a statistical analysis of random fatigue. Although the information it provides is unsatisfactory in some respects, the damage rate is certainly not irrelevant as a sensible criterion of fatigue failure. Consequently, it is appropriate to examine the statistical behavior of the damage rate (which may be used as a simple index of product reliability) if we want to avoid the serious penalty imposed by a more comprehensive analysis.

It is possible to obtain a rather complete statistical description of the damage rate by means of eq.(7) from suitable information about the stress variation. It is expedient to augment the damage rate R by another stochastic function u as follows:

$$R = \frac{\beta|S|^{\beta-1}|\dot{S}|}{4\lambda} \geq 0 \text{ and } u = |\dot{S}| \geq 0 \qquad (22)$$

The joint probability density r of the stochastic functions R and u is related to the joint probability density q of the random variables S and $\dot{S}$ as follows:

$$r(R,u;t) = \frac{q(S,\dot{S};t)}{\left|J\left(\frac{R,u}{S,\dot{S}}\right)\right|} \qquad (23)$$

where the Jacobian J of the transformation is given by:

$$J\left(\frac{R,\ u}{S,\ \dot{S}}\right) = \begin{vmatrix} \frac{\partial R}{\partial S} & \frac{\partial R}{\partial \dot{S}} \\ \frac{\partial u}{\partial S} & \frac{\partial u}{\partial \dot{S}} \end{vmatrix} = \pm \frac{\partial R}{\partial S} \tag{24}$$

which leads to the following result:

$$|J| = \left|\frac{\partial R}{\partial S}\right| = (\beta-1)\ R \left(\frac{\beta u}{4\lambda R}\right)^{\frac{1}{\beta-1}} \tag{25}$$

A substitution then yields:

$$r(R,\ u;t) = \frac{F(R,\ u,t)}{(\beta-1)R}\left(\frac{4\lambda R}{\beta u}\right)^{\frac{1}{\beta-1}} \tag{26}$$

where the function F is given by:

$$F(R,\ u,\ t) = q\left[\left(\frac{4\lambda R}{\beta u}\right)^{\frac{1}{\beta-1}},\ u;t\right] + q\left[-\left(\frac{4\lambda R}{\beta u}\right)^{\frac{1}{\beta-1}},\ u;t\right]$$

$$+\ q\left[\left(\frac{4\lambda R}{\beta u}\right)^{\frac{1}{\beta-1}},-u;t\right] + q\left[-\left(\frac{4\lambda R}{\beta u}\right)^{\frac{1}{\beta-1}},\ -\ u;t\right] \tag{27}$$

The probability density function p of the damage rate R is then given by the following marginal probability density function:

$$p(R,\ t) = \int_{-\infty}^{+\infty} r(R,\ u;t)\ du =$$

$$\frac{1}{(\beta-1)R}\int_0^{\infty}\left(\frac{4\lambda R}{\beta u}\right)^{\frac{1}{\beta-1}} F(R,\ u,\ t)\ du \tag{28}$$

This relation is valid for a general probability distribution. In the event of a stationary normal stress variation about a zero mean stress level, the joint probability density function q of S and $\dot{S}$ is given by:

$$q(S,\ \dot{S}) = \frac{1}{2\pi\sigma_S\sigma_{\dot{S}}}\exp\left[-\left(\frac{S^2}{2\sigma_S^2} + \frac{\dot{S}^2}{2\sigma_{\dot{S}}^2}\right)\right] \tag{29}$$

A substitution then yields:

$$F(R,\ u) = \frac{2}{\pi\sigma_S\sigma_{\dot{S}}}\exp\left[-\frac{1}{2\sigma_S^2}\left(\frac{4\lambda R}{\beta u}\right)^{\frac{2}{\beta-1}}\right]\exp\left(-\frac{u^2}{2\sigma_{\dot{S}}^2}\right) \tag{30}$$

The probability density function of the damage rate is accordingly given by:

$$p(R) = \frac{2}{\pi(\beta-1)\sigma_S\sigma_{\dot{S}}R}\int_0^{\infty}\left(\frac{4\lambda R}{\beta u}\right)^{\frac{1}{\beta-1}}$$

$$\exp\left[-\frac{1}{2\sigma_S^2}\left(\frac{4\lambda R}{\beta u}\right)^{\frac{2}{\beta-1}}\right]\exp\left(-\frac{u^2}{2\sigma_{\dot{S}}^2}\right)du \tag{31}$$

The statistical moments μ_n are then calculated as follows:

$$\mu_n = \int_0^{\infty} R^n\ p(R)\ dR$$

$$= \frac{2}{\pi(\beta-1)\sigma_S\sigma_{\dot{S}}}\int_0^{\infty}\int_0^{\infty} R^{n-1}\left(\frac{4\lambda R}{\beta u}\right)^{\frac{1}{\beta-1}}$$

$$\exp\left[-\frac{1}{2\sigma_S^2}\left(\frac{4\lambda R}{\beta u}\right)^{\frac{2}{\beta-1}}\right]\exp\left(\frac{u^2}{2\sigma_{\dot{S}}^2}\right)dR\ du \tag{32}$$

Evaluation of this double integral yields the following expression for the statistical moments of arbitrary order:

$$\mu_n = \frac{1}{\pi}\left(\frac{\beta\Omega_S 2^{\frac{\beta}{2}}\sigma_S{}^{\beta}}{4\lambda}\right)^n \Gamma\left(\frac{n+1}{2}\right)\Gamma\left[\frac{n(\beta-1)+1}{2}\right]$$

$$\text{with } \Omega_S = \frac{\sigma_{\dot{S}}}{\sigma_S} \tag{33}$$

where the gamma function is given by:

$$\Gamma\left(\frac{n+1}{2}\right) = \Gamma\left(\frac{2m+1}{2}\right) = \Gamma(m + \tfrac{1}{2}) =$$

$$\frac{(2m)!}{m!\,2^{2m}}\sqrt{\pi}\ \text{when } n = 2m \tag{34}$$

while it is given by:

$$\Gamma\left(\frac{n+1}{2}\right) = \Gamma\left(\frac{2m+2}{2}\right) = \Gamma(m + 1) = m!\quad \text{when } n = 2m + 1 \tag{35}$$

with m = 0,1,2, ... since n is an integer.

As usual, the zero-order moment is equal to unity:

$$\mu_0 = 1 \tag{36}$$

The ensemble mean m of the damage rate is given by the first-order moment:

$$m = \mu_1 = \frac{\beta\Omega_S}{4\pi\lambda}\ 2^{\frac{\beta}{2}}\ \sigma_S{}^{\beta}\ \Gamma\left(\frac{\beta}{2}\right) \tag{37}$$

while the ensemble mean square s is given by the second-order moment as follows:

$$s = \mu_2 = \frac{1}{2\sqrt{\pi}}\left(\frac{\beta\Omega_S}{4\lambda}\ 2^{\frac{\beta}{2}}\ \sigma_S{}^{\beta}\right)^2\Gamma\left(\beta - \frac{1}{2}\right) =$$

$$\frac{\pi^{\frac{3}{2}}\Gamma\left(\beta - \frac{1}{2}\right)}{2\Gamma^2\left(\frac{\beta}{2}\right)}\ m^2 \tag{38}$$

The standard deviation σ of the damage rate is accordingly given by:

$$\sigma = \sqrt{s-m^2} = m\sqrt{\frac{\pi^{\frac{3}{2}}\,\Gamma\left(\beta - \frac{1}{2}\right)}{2\Gamma^2\left(\frac{\beta}{2}\right)} - 1} = k(\beta)m \qquad (39)$$

Consequently, the standard deviation is proportional to the mean value of the damage rate, and the ratio k between them depends only upon the fatigue exponent β.

The statistical moments and the probability distribution of the damage rate depend upon the material parameters β and λ, as well as the mean frequency Ω_S and standard deviation σ_S of the random stress variation. The statistical parameters Ω_S and σ_S may be evaluated from the spectral density $P_{SS}(\omega)$ of the stress variation as follows:

$$\Omega_S = \frac{\sigma_{\dot{S}}}{\sigma_S} \text{ with } \sigma_S^2 = \frac{1}{\pi}\int_0^\infty P_{SS}(\omega)\,d\omega$$

$$\text{and } \sigma_{\dot{S}}^2 = \frac{1}{\pi}\int_0^\infty \omega^2 P_{SS}(\omega)\,d\omega \qquad (40)$$

<u>Random Vibration</u>

The statistical moments μ_n of the damage rate R indicated by eq. (33) depend upon the mean frequency Ω_S and standard deviation σ_S of the random stress variation. These statistical parameters are related, in turn, by eq. (40) to the spectral density $P_{SS}(\omega)$ of the stress variation S, which depends upon the corresponding spectral density $P_{FF}(\omega)$ of the stationary random excitation F as follows:

$$P_{SS}(\omega) = |H_{SF}(\omega)|^2\, P_{FF}(\omega) \qquad (41)$$

where $H_{SF}(\omega)$ is the frequency response or system function of the stress (output) with respect to the excitation (input) of a time-invariant linear system. The standard deviation σ_S and $\sigma_{\dot{S}}$ of S and $\dot{S}$, respectively, are then given by:

$$\sigma_S^2 = \frac{1}{\pi}\int_0^\infty |H_{SF}(\omega)|^2\, P_{FF}(\omega)\,d\omega$$

$$\text{and } \sigma_{\dot{S}}^2 = \frac{1}{\pi}\int_0^\infty \omega^2 |H_{SF}(\omega)|^2 P_{FF}(\omega)\,d\omega \qquad (42)$$

Consider the situation depicted in Fig. 6. Let us assume that the spring is the critical element in the simple mechanical system, which also consists of a parallel dashpot with a mass element exposed to a random force variation. The differential equation of motion is accordingly given by:

$$m\ddot{x} + c\dot{x} + kx = F(t) \qquad (43)$$

where x is the displacement of the mass m induced by the force F, k is the elastic spring constant, and c is the viscous damping coefficient. The stress S in the damage-sensitive element is proportional to the displacement x as indicated below:

$$S = \gamma x \qquad (44)$$

The system function is then given by:

$$H_{SF}(\omega) = \frac{S}{F} = \frac{\gamma}{k-m\omega^2+jc\omega} \text{ with } F = F_0 e^{j\omega t}$$

$$\text{and } S = S_0 e^{j\omega t} \qquad (45)$$

which is equivalent to:

$$H_{SF}(\omega) = \frac{\alpha}{1-\rho^2+j2\xi\rho} \text{ with } \alpha = \frac{\gamma}{k},\ \rho = \frac{\omega}{p},$$

$$\xi = \frac{c}{2mp},\ p = \sqrt{\frac{k}{m}} \qquad (46)$$

where p is the natural frequency, and ξ is the damping ratio of the mechanical system. It follows that:

$$|H_{SF}(\omega)|^2 = \frac{\alpha^2}{(1-\rho^2)^2+(2\xi\rho)^2} \qquad (47)$$

Let us assume that the random excitation is a "white noise" with a uniform spectral density as indicated below:

$$P_{FF}(\omega) = P = \text{constant} \qquad (48)$$

A substitution of eq. (47) and (48) into (42) then yields:

$$\sigma_S^2 = \frac{\alpha^2 pP}{\pi}\int_0^\infty \frac{d\rho}{(1-\rho^2)^2+(2\xi\rho)^2} \text{ and } \sigma_{\dot{S}}^2 =$$

$$\frac{\alpha^2 p^3 P}{\pi}\int_0^\infty \frac{\rho^2\,d\rho}{(1-\rho^2)^2+(2\xi\rho)^2} \qquad (49)$$

Evaluation of these integrals then yields the following results:

$$\sigma_S = \alpha\sqrt{\frac{pP}{4\xi}} \text{ and } \sigma_{\dot{S}} = \alpha p\sqrt{\frac{pP}{4\xi}} \text{ with } \Omega_S = \frac{\sigma_{\dot{S}}}{\sigma_S} = p \qquad (50)$$

and the mean frequency of the stress variation coincides with the natural frequency of the mechanical system irrespective of the damping ratio. A substitution of eq. (50) into eq. (33) then yields the following expression for the statistical moments:

$$\mu_n = \frac{1}{\pi}\left[\frac{\beta p\alpha^\beta}{4\lambda}\left(\frac{pP}{2\xi}\right)^{\frac{\beta}{2}}\right]^n \Gamma\left(\frac{n+1}{2}\right)\Gamma\left[\frac{n(\beta-1)+1}{2}\right] \qquad (51)$$

which depends upon the material parameters β and λ, the elastic constant α, the system parameters p and ξ, and the spectral density P of the random excitation. This result is contingent upon a variety of assumptions as indicated below:

1) material: cycle-dependent, linear damage accumulation with a power law of failure.

2) system: time-invariant, linear system with a single degree of freedom.

3) load: stationary, normal excitation with a uniform spectral density.

The statistical variability of the material is ignored in this analysis, which refers to the low-level, high-cycle type of fatigue failure. It is possible to relax some of these assumptions to account for a constant stress level, several degrees of freedom, or variable spectral density by a suitable extension of the above procedure.

Conclusion

It is difficult to evaluate product reliability from a statistical description of the random excitation due to a serious impasse at a certain intermediate stage of the analysis. Once we appreciate the fundamental nature of this difficulty, however, it is possible to devise a realistic strategy along the practical lines indicated below:

(1) We may terminate the mathematical analysis at this point and use the symbolic expression for the damage rate as a simple index of fatigue behavior.

(2) We may dispose of this obstacle by a numerical analysis of sample data to obtain more relevant information about product reliability.

These options are not mutually exclusive, and it may be advisable to combine them to some extent. The former may be used to compare alternative design proposals, while the latter may be used to evaluate hardware performance requirements. In either instance, a computer program may be devised to expedite the analysis.

The statistical chain that connects the product reliability to the random vibration contains a weak link that is not strong enough to support a complete analysis. However, we may start at either end of this chain and move toward the crucial link from opposite directions as indicated below:

1) Random vibration: The spectral density $P_{FF}(\omega)$ of the random excitation F(t) may be used to determine the mean frequency Ω_S and standard deviation σ_S of the stress variation S(t) from the frequency response $H_{SF}(\omega)$ of the mechanical system. This information may be used to calculate the statistical moments $\mu_n(R)$ of the damage rate R(t) from the material parameters β and λ as indicated by the schematic diagram of Fig. 7 and the expressions indicated below:

$$\sigma_S = \sqrt{\frac{1}{\pi}\int_0^\infty P_{FF}(\omega)\,|H_{SF}(\omega)|^2 d\omega} \tag{52}$$

$$\sigma_{\dot{S}} = \sqrt{\frac{1}{\pi}\int_0^\infty \omega^2 P_{FF}(\omega)\,|H_{SF}(\omega)|^2 d\omega} \tag{53}$$

$$\Omega_S = \frac{\sigma_{\dot{S}}}{\sigma_S} \tag{54}$$

$$\mu_n = \frac{1}{\pi}\left(\frac{\beta\Omega_S 2^{\frac{\beta}{2}}\sigma_S^{\beta}}{4\lambda}\right)^n \Gamma\left(\frac{n+1}{2}\right)\Gamma\left[\frac{n(\beta-1)+1}{2}\right] \tag{55}$$

This result is based on the following relation between the damage rate and the stress variation:

$$R(t) = \frac{\beta}{4\lambda}\,|S(t)|^{\beta-1}|\dot{S}(t)| = \frac{1}{4\lambda}\left|\frac{d}{dt}|S(t)|^{\beta}\right| \tag{56}$$

It is necessary to modify the damage relation indicated above to account for more complicated stress conditions. This aspect of fatigue behavior is a good topic for basic research in engineering mechanics.

2) Product reliability: It is possible to evaluate product reliability in terms of the probability of survival $\rho(t)$ to time t from the probability distribution P(L) of the fatigue life L associated with a random excitation. This information may be derived from the probability distribution Q(D,t) of the cumulative damage D, which is given by integration of the probability density q(D,t) as indicated by the schematic diagram of Fig. 8 and the relations described below:

$$\rho(t) = 1 - P(t) \tag{57}$$

$$P(L) = 1 - Q(1,L) \tag{58}$$

$$Q(1,t) = \int_0^1 q(D,t)\,dD \tag{59}$$

The cumulative damage D(t) is given by the time integral of the damage rate R(t) as follows:

$$D(t) = \int_0^t R(t)\,dt \tag{60}$$

Despite the simple appearance of this elementary relation, the statistical connection between these random variables is difficult to establish. This is a good topic for basic study in applied mathematics.

The application of a computer program to overcome the deficiency indicated above would require ensemble records of the random excitation to provide sample data for a numerical analysis. The individual data records would permit a determination of the corresponding stress variation from the impulsive response of the mechanical system. This would provide the information required for a computer program to evaluate product reliability.

The integrity of the results would depend upon a variety of considerations, which include the statistical nature of the random excitation, the response characteristic of the mechanical system, and the fatigue behavior of the material. Among the potential sources of error is the possibility of a statistical interaction between random load and material parameters, the distortion produced by non-linear components in the mechanical system, inelastic strain variations due to acute stress concentrations, and indefinite residual or biaxial stress conditions.

The math model also constitutes a source of error due to the approximate nature of the linear damage hypothesis and the power law of failure. In this respect, there is a curious tendency to indict the math model (especially the linear damage hypothesis) for every discrepancy between theoretical calculations and experimental observations without a critical examination of other factors such as those indicated above. A significant part of this error (which is not incompatible with the typical scatter of fatigue data) could be due to a statistical bias associated with a random variation of the material parameters. However, even if it were entirely responsible for this discrepancy, the math model would continue to enjoy wide popularity in view of its attractive simplicity and practical utility.

After more than a century of empirical study, fatigue data is now available in sufficient quantity to satisfy even the most hungry appetite. The reward for such a diet, however, is a very unhappy state of confusion. The only remedy for this headache is a more sensible attitude toward specific deviations and a constructive desire to investigate the general pattern of material behavior. While this may not eliminate the source of the discomfort, it should alleviate the headache to some degree and facilitate a logical analysis of fatigue failure and product reliability due to random vibration.

References

[1] O. H. Basquin, "The exponential law of endurance tests," Amer. Soc. Test. Mater. Proc., vol. 10, pp. 625-630, 1910.

[2] M. A. Miner, Estimation of Fatigue Life with Particular Emphasis on Cumulative Damage, Metal Fatigue. New York: McGraw-Hill, 1959, pp. 278-289.

[3] W. H. Raser, Jr., Fatigue Analysis and Measurement of Random Loading, Metal Fatigue: Theory and Design, New York: Wiley, 1969, pp. 247-288.

[4] A. Sorensen, Jr., "A general theory of fatigue damage accumulation," Trans. ASME, vol. 91, pp. 1-14, Mar. 1969.

[5] A. Powell, "On the fatigue failure of structures due to vibrations excited by random pressure fields," J. Acoust. Soc. Amer., vol. 30, pp. 1130-1135, Dec. 1958.

[6] A. Sorensen, Jr., "Product reliability and random fatigue," Trans. IEEE, vol. R-20, pp. 244-248, Nov. 1971.

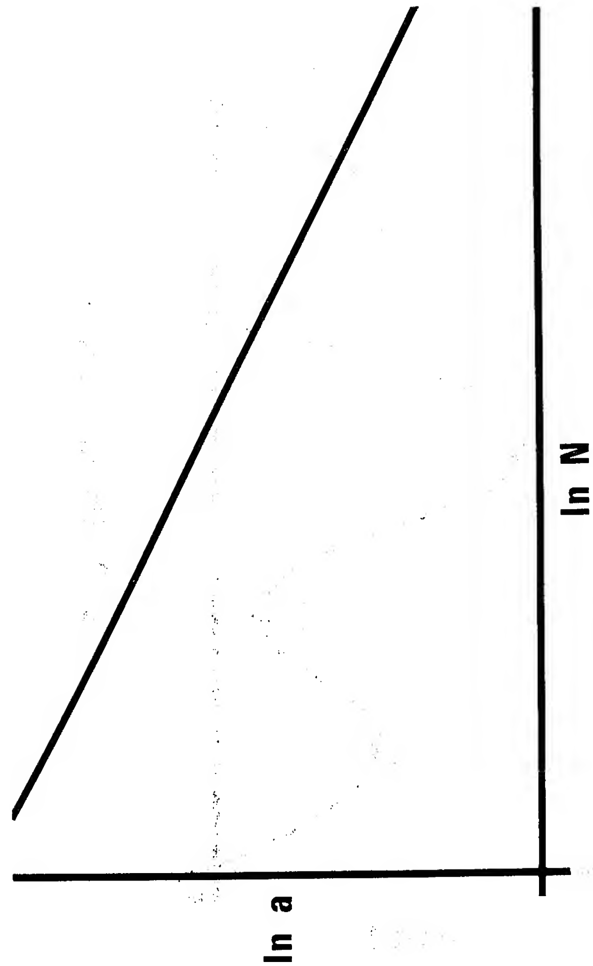


Fig. 1 Fatigue Life

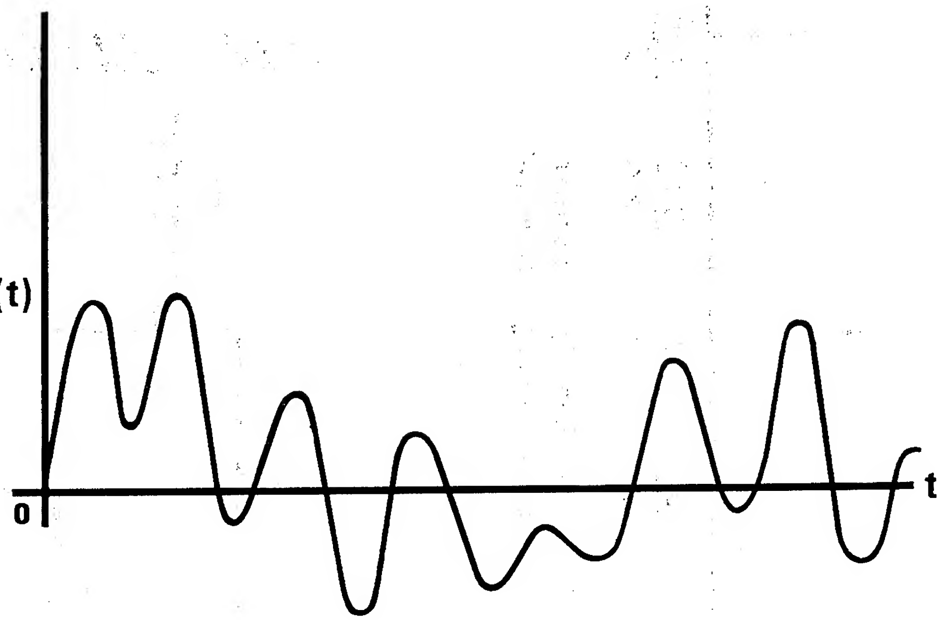


Fig. 2 Irregular Waveform

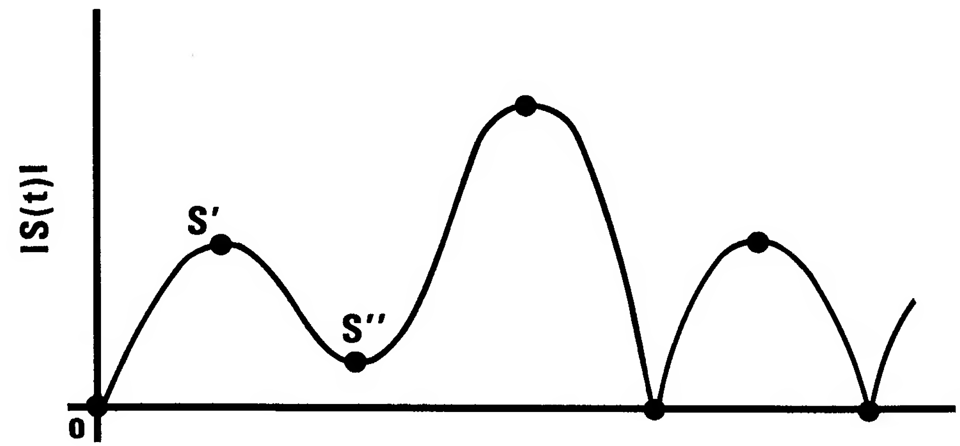


Fig. 3 Stress Profile

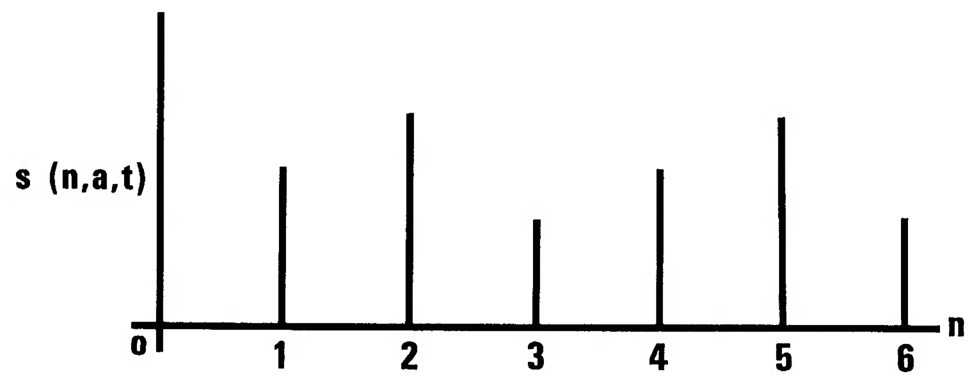


Fig. 4 Probability of Stress Maximum or Minimum

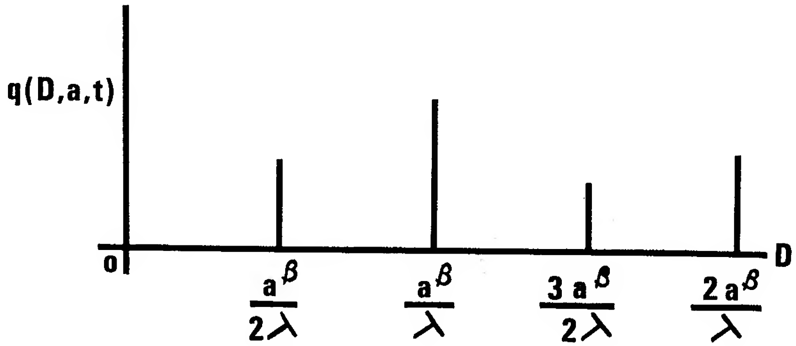


Fig. 5 Probability of Positive or Negative Damage Increment

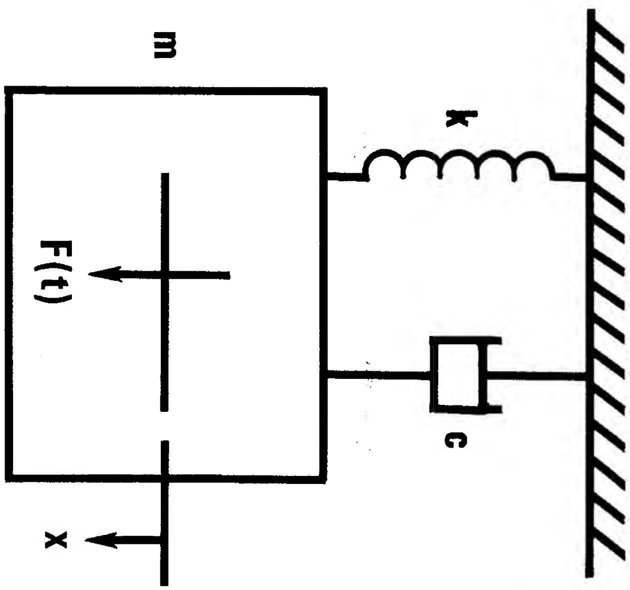


Fig. 6 Mechanical System

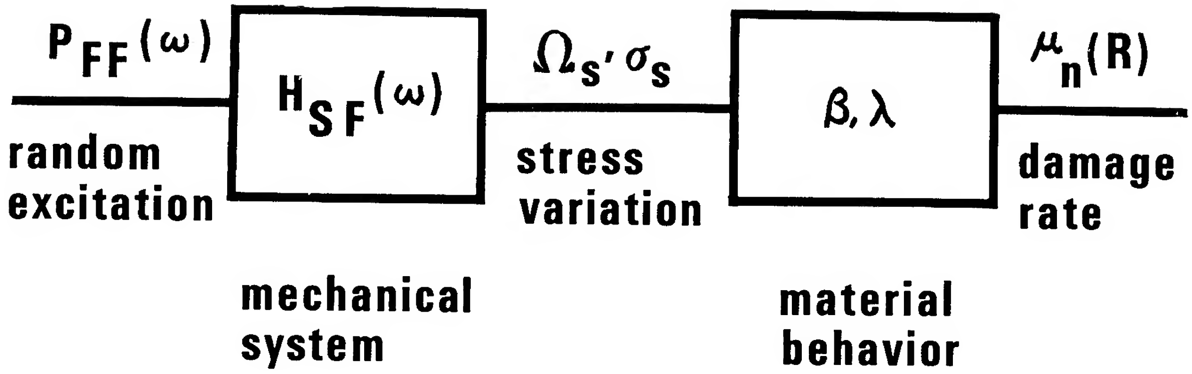


Fig. 7 Random Vibration

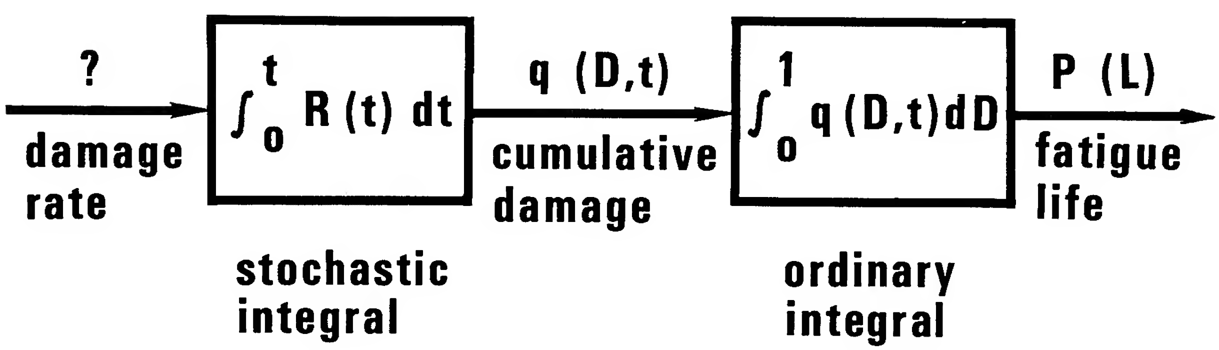


Fig. 8 Product Reliability

MARKETING RELIABILITY PROGRAMS IN THE 70's

John L. Hesse

SAFEGUARD System Evaluation Agency
Kwajalein Field Office
Kwajalein, Marshall Islands

Summary

A methodology is presented to quantify the costs of changing system availability and reliability in terms of incremental changes in system performance capability. Two simulations of the system are described, the first of which relates assurance data to subsystem-level failure, downtime, and repair distributions. The second simulation relates these data to system performance capability descriptors in a manner suitable for use in management decision making. Validity and output analysis are discussed.

Introduction

Although reliability engineering has matured as a discipline in the past decade under the pressure of increasing user and manufacturing requirements, concern continues within the discipline about the appropriate method for presenting results to high-level managers. The existing literature has concentrated on increasingly sophisticated availability and reliability programs together with new statistical and analytical approaches to failure analysis, test design, and criticality determination, leaving the marketing aspects of program management essentially untouched.[1,2,3]

Prior to, during, and after the creation of a complex system, corporate managers normally expect a reliability group to identify, and recommend corrective action for, those factors in system design which may cause degradation in operating performance from an availability or reliability standpoint. Although design specifications may exist which define satisfactory performance for mildly innovative developmental efforts, research and development efforts on new systems which cause extensions to be sought to state-of-the-art technology often rely on design goals. Definitions of satisfactory performance then become subjective, and the costs of upgrading availability or reliability so as to increase the level of system performance must be traded off against such intangibles as user satisfaction and developer reputation. Such tradeoff decisions require derivation of relationships between performance parameters and availability-reliability factors.[4] Concisely stated, we seek a way to make reliability engineering results meaningful to cost conscious, performance-minded managers. The methodology to follow has proven successful in solving this problem.

Data Available

Any novel methodology must reflect real-world data availability if it is to find acceptance and use. Data on a system evolving from conception to fielding may come from several sources, as follows:

System Performance and Operating Environment

Design specifications and user requirements
System analysis calculations
Prototype testing
Historical analogies

Reliability Engineering Data

Stress analysis
Prototype testing
Piecepart data bases such as FARADA and RADC II
Historical analogy data
Configuration documentation
Component testing
Prototype operating time records
Prototype maintenance reports
Failure analysis reports
Spares lists
Human interface data

System performance data may be highly theoretical, as in the case of systems where full prototypes are too expensive to construct, or where the expected operating environment cannot be reproduced or fully simulated in a test situation. An example of this extreme was found in NASA's initial Lunar Rover vehicle. At the other extreme, motor car manufacturers exhaustively test new models before release to the public.[5] Computer simulation models are often used to bridge testing gaps in determining bounds on system performance by interpolation between data points on the system capability envelope.

Although reliability engineers normally have a large historical experience base from which to draw estimates of failure and repair times, it will frequently be found that the necessary data are not available, particularly for new parts and parts peculiar to a specific application. Test programs and detailed laboratory failure analyses, together with sophisticated statistical treatment of results, may still be required to supplement available information.[6]

A larger problem may arise in configuration definition, which may be highly time-variant during the development and initial use phases. This problem is, in part, caused by assurance technology feedback to design groups in the form of part improvement programs, circuit and equipment redesign, inclusion of redundancy, and reallocation of requirements. Such action may or may not be cost effective in the larger system sense.[7] Human interfaces and spares policy are particularly difficult to define, since their determination may be beyond developer control in any real sense. Their influence on overall system performance may, however, overshadow all other considerations combined, and hence should be carefully considered.

The Availability-Reliability Model

Let us assume that we are concerned with a large evolving system which approaches state-of-the art performance and complexity. Since our data, as previously noted, is mainly available at the piecepart level, we will forego hand calculations in favor of an availability-reliability model designed for a digital computer. In nonspecific terms, we wish to construct a model sufficiently general to accept subsystem configuration data as input, together with failure, repair, and replacement data for each subsystem element defined. These latter data elements should, ideally, be probabilistic in nature, and should be entered into our model

as density functions. In order to clarify our terminology, we will define replacement as an element for element exchange situation which maximizes system uptime. Repair then denotes an on-line fault identification, diagnostic, and corrective maintenance sequence resulting in potentially lengthy system downtime. In this paper, however, we will use the term repair to encompass both repair and replacement situations. The form of our data thus implies a non-deterministic, or Monte-Carlo, type of model which would be usable for a succession of systems requiring analysis.

Application of the model begins by initiating one of a number of random walks. Failure and repair density functions are randomly sampled, resulting in the acquisition of explicit failure and repair times. Once the subsystem element and its failure and repair times have been identified within the model, the effect at higher levels is computed by chaining through the connectivity diagram. The resulting subsystem status changes are saved, analyzed statistically, and employed in producing subsystem failure and repair distributions. Once a random walk has been completed and the subsystem status changes recorded, a new random walk will be initiated. Random walks continue until a pre-determined total of subsystem downs (subsystem inoperative) has been obtained. This pre-determined total is based on desired output statistics, and is input data to the model.

After all random walks have been concluded, the walk histories are retrieved and analyzed. Histograms are now constructed for subsystem downtime, time between subsystem failures, and the time to repair the subsystem. Separate listings can be prepared of elements contributing significantly to subsystem downtime (mission-essential items). These items may be isolated by means of a statistical analysis for outlying observations. Subsystem availability, downtime, and time-to-failure distributions are now constructed from generated histograms. Let us consider these distributions as an operating profile for our subsystem, and set them aside for later use (for each subsystem in the system).

The System Model

We stated earlier that our overall objective is to translate reliability engineering results into a form meaningful to cost and performance conscious managers. At system performance capability level, then, we wish to relate the operating profiles calculated in our availability-reliability model to some set of quantifiable descriptors of system performance capability that is meaningful to management. To restate this, we seek to define availability-reliability leverage in units of system performance capability. An example of such leverage, in ballistic missile defense analysis, might lie in relating interceptor availability to the number of MINUTEMAN missiles preserved against a given hostile scenario. Given the value in dollars of each MINUTEMAN, together with the cost in dollars of incrementing interceptor availability, the leverage thus defined can be evaluated as a dollar tradeoff between interceptors and defended MINUTEMAN missiles. In this fashion, interceptor availability calculations assume meaning to managers. This example has been somewhat oversimplified in order to illustrate the point.

Let us consider the construction of a model of system operational capability. Three significant items of data are available to us in constructing such a model. These include intended modes of system operation, the operating environment, and our subsystem operating profiles. Our model, then, might be a functional simulation of the system which represents the performance capabilities of each included subsystem, within the operating environment for each subsystem. Expanding our previous example, we might model the search volume of a target and interceptor tracking radar, in a nuclear blackout environment, as defined on a pulse-by-pulse basis. To this we might add a representation of interceptor flight dynamics, in various nuclear blast regimes, as a function of radar-issued discrete steering commands. Operating rules by which the radar and interceptor perform together to intercept a hostile reentry vehicle must, of course, be superimposed on the subsystem representations, as must the realities defined by the previously calculated subsystem operating profiles.

Certain required general characteristics of our system model can be established. Since the operating profile is input in the form of subsystem availability, downtime, and time-to-failure distributions, sampling must take place from these distributions on a random basis to obtain discrete values for calculation. The structure of the subsystem models, which with operating rules work together to comprise the system model, may be either time or event oriented. If an event-oriented model is chosen, calculations advance by steps whose time size varies dependent on events which occur. System downs are an example of events in which we might have an interest.

In our discussion of system operation, we should include the possibility of degraded operation due to non-catastrophic part failures. If acceptable operation is relatively binary in nature and does not encompass performance degradation below rigid specifications, degraded operation need not be considered. Otherwise, part failure must be related to an effect on a system capability descriptor. This must be separately determined by constructing appropriate histograms for this type of part using the availability-reliability model, fitting distributions, and adding input to the system performance capability model to account for the cause and effect relationships. Model complexity thus increases significantly.

The output, or post-processing, section of our system model must generate the relationship we seek between reliability engineering inputs, or our operating profiles, and the system capability descriptors we have chosen as leverage identifiers on overall performance. Inputs to post-processing may be any or all of the following:

Subsystem response opportunity
Subsystem response
Performance duration
Reason for subsystem response
System response per subsystem response

The correlations we seek between cause and effect, for example the effect of failure history of a given subsystem on overall system performance, are now quantitatively available. Applying a generalization to the term subsystem, we can see the potential for relating element performance at any level to an incremental cost in system-level performance.[8]

Completing our example drawn from ballistic missile defense analysis, we may select MINUTEMAN missiles saved, against a particular hostile scenario, as our system capability descriptor. Each missile saved has a readily determined dollar value, and we may choose to use these funds to increase the availability or reliability of our interceptors. Since our system performance capability model relates these factors, in terms of our operating profile, to intercepts lethally performed against hostile reentry vehicles and thus to MINUTEMAN missiles preserved against the hostile threat, a replotting of incremental MINUTEMAN missiles saved versus incremental interceptor availability or reliability required to achieve the saving allows managers to trade dollars off. We are now in a position to understand how much availability and reliability is worth purchasing in units of

system performance.

Validity

We have discussed two closely-linked simulation models which, in concert, will operate to relate reliability engineering data at element level to system performance capability. A few words about model validation are in order at this point. If managers are to base potentially expensive and program-significant decisions on our work, the assurance in decision model validity must be as high as possible. System prototype availability is a distinct plus in model validation if tight configuration control exists to allow close definition of production versus prototype differences. Availability of prototypes at the subsystem level is useful, but their exclusive use leaves operating rules by which subsystems act together unverified. The fact that a system operational capability model can be constructed and run will assist in defining these operating rules even if validated only at the subsystem level. We should note that it may not be possible to introduce test feedback for model validation, at any level, until rather late in the development cycle, particularly for military systems where concurrent development and deployment are planned. The analyst must, in cases like this, take great care in the confidence he attributes to his results until test data do become available.

Summary and Conclusions

We sought some method whereby assurance technology could exert its proper influence on managers better versed in costs and performance than in reliability engineering. The solution we have arrived at is to develop, and if possible validate, two disjoint but closely related models. The first model accepts reliability and availability data at subsystem element level together with subsystem configuration information. It chains the data through the configuration to develop subsystem-level estimates of reliability and availability. These results, for all subsystems, then drive a system performance capability simulation which produces descriptors of the leverage exerted by assurance factors on system performance. Figure 1 illustrates this scheme. As in any mathematical modeling technique, we must apply good judgement to the problem of assessing output validity based on data quantity and quality as well as configuration management principles.

The methodology described has been successfully tested with a large military system development effort. Managers to whom results have been presented have expressed satisfaction in the understandable quantification obtainable, and the reliability groups involved have been better able to assess their program revision proposals prior to presentation to management. In an era of cost and performance consciousness, then, the marketability of assurance technology is significantly enhanced by use of the techniques presented here.

Bibliography

1. Muench, J. O., "A Complete Reliability Program," Proc. 1972 Ann. Rel. and Maint. Symposium, San Francisco California, January 25-27, 1972, pp. 20-23.

2. Price, D. C., "Testing of Mechanisms with Small Sample Sizes," Proc. 1969 Ann. Symposium on Reliability, Chicago, Illinois, January 21-23, 1969, pp. 562-566.

3. Miller, R. N., "Decision Theory in Reliability and Project Management," Proc. 1971 Ann. Symposium on Rel., Washington, D. C., January 12-14, 1971, pp. 376-382.

4. Gnedenko, B. V., Belyayev, Y. K., and Solovyev, A. D., Mathematical Methods of Reliability Theory, Academic Press, New York, New York, 1969, pp. 77-78.

5. Lloyd, D. K. and Lipow, M., Reliability: Management, Methods, and Mathematics, Prentice-Hall, Englewood Cliffs, N. J., 1962, pp. 37-58.

6. Shooman, M. L., Probabilistic Reliability: An Engineering Approach, McGraw-Hill Book Company, New York, New York, 1968, pp. 160-180.

7. Von Alven, W. H., Reliability Engineering, Prentice-Hall, Englewood Cliffs, N. J., 1964, pp. 268-278.

8. Rau, J. G., Optimization and Probability in Systems Engineering, Van Nostrand Reinhold Company, New York, New York, 1970, pp. 11-16.

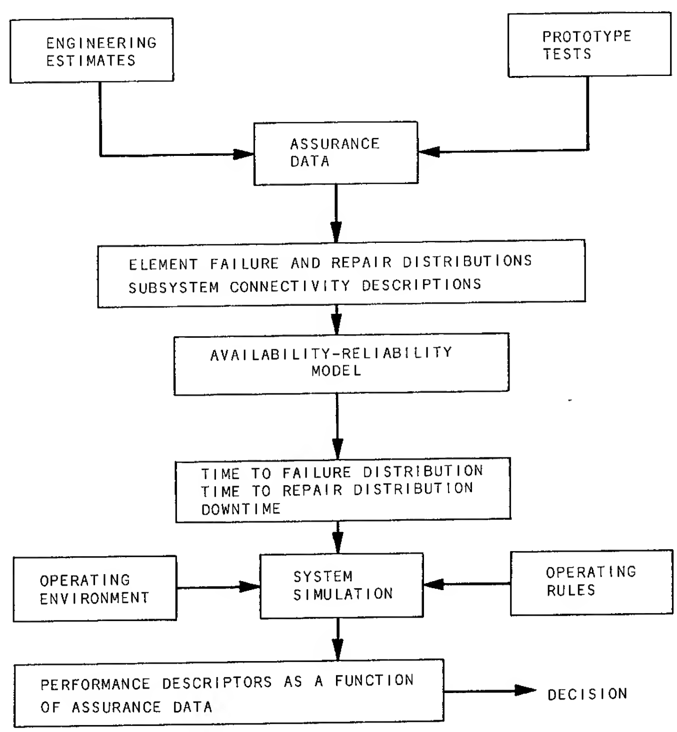


FIGURE 1

OVERALL METHODOLOGY

SAFETY IN POSEIDON TECHNICAL MANUALS

Walter F. Lutzweit
Missile Systems Division Product Assurance
Lockheed Missiles and Space Company
Sunnyvale, California

Introduction

The Poseidon Missile System is one of several subsystems of the Poseidon Fleet Ballistic Missile (FBM) Weapon System. Other subsystems are fire control, navigation, communications, launcher and the ship submersible, ballistic, nuclear (SSBN). The missile subsystem includes instrumentation for converting to evaluation configurations, and special vehicles required for training of FBM personnel and for SSBN shipyard checkout. The missile and its equipment are supported by other elements of the weapon system, and its support elements include a documentation system. The documentation system was developed to provide orientation, training, and maintenance of equipment and to ensure that the missiles are safe, reliable and efficient.

This paper concerns the safety review of the documentation system. It is a first-hand report. The work was done by experts in the Reliability and Product Safety Department of the Lockheed Missile Systems Division, Product Assurance organization. The System Safety Specialists directing the effort were not concerned with whether or not the problem described in the manual was one of plant or product, hygiene or medical, or system. Primary objectives were to warn of, and minimize, hazards from foreseeable events described in the manuals. A guiding principle was the conservation of life and resources in all evaluation as well as assuring satisfaction of safety requirements imposed by contract.

Highlights of problems and safety accomplishments are presented for system, shorebase and fleet equipment manuals. Conclusions support a position that experience and acumen are more desirable traits than quantitative safety skills on a program of this scope. The compendium[1] referenced in this report will prove a valuable tool for those reviewing, or responsible for, the inclusion of safety discipline in documents. Major functions and interfaces are depicted in Figure 1.

Although the last word will not be written about Poseidon Technical Manuals for many years to come, it would seem appropriate to mention that the customer did not cite the omission of a single safety item from any of the original publications.

Task Description and Safety Considerations

The Poseidon Missile Technical Manuals can be classified into ten areas of responsibility, as shown in Figure 2. However, this illustration does not fully convey the volume of documentation that exists. For a better appreciation of the magnitude and complexity of this element of weapon system support, a break-down of the small block identified as "Missile Surface Support Equipment Manuals" in Figure 2 may be used. Utilizing such a break-down, one finds there are one hundred and forty-five different equipment manuals under this sub-grouping. Generally, each of these covered description, operation, and maintenance of hardware used aboard SSBN, Tender, and shorebase facilities. Hardware complexity ranged from Poseidon missile component containers to the truck and rail van (TARVAN) system. All of the manuals were reviewed. The stated purposes of the safety review were to assure that personnel safety and equipment damage were considered, and the necessary precautionary measures were incorporated in the technical manual. The Reliability and Product Safety Review included, but was not limited to, the following aspects:

1. Inclusion and effective use of warning and caution notices.
2. Inclusion of a Safety Summary.
3. Proper and safe sequence of operations within the individual work sections.
4. Incorporation of general safety requirements and principles from a Government Safety Document List (reference compendium[1]).
5. Recognition of hazard-producing factors identified in the Lockheed Missiles and Space Company Safety and Industrial Hygiene Standards.
6. Assurance that limited life aspects were considered.
7. Verification of accomplishment of critical safety steps.
8. Assurance of clarity of instruction.

Real-world safety criteria are indispensable for an engineer performing a safety review if he is expected to consider all potential hazards. Commercial standards for safety of devices, systems, and material are as close as the nearest Underwriters Laboratory. Safety standards for plant, some product, and medical are available in a single source from the recently published Occupational Safety and Health Standards (29 May 1971). Unfortunately for the defense industries, the availability of safety criteria from a single source is as much an illusion as the rainbow's pot of gold.

The product mix for the Poseidon missile system is such that safety criteria for plant, product, industrial hygiene, and system had to be available for use by the safety reviewer. The problem was solved by compiling a compendium[1] of Military Specifications, Standards and Publications, Department of Defense and Department of Transportation, and Bureau of Explosives of the American Association of Railroads rules and regulations pertaining to safety. This list was then integrated with unique Poseidon program safety requirements. The next step was to obtain the documents having an impact on system or equipment manuals in order to achieve a solid base of safety criteria against which to review the manuals. An experienced engineer could then make worthwhile judgements as to how safety was considered in the manual and whether any safety principles were abused.

Additional benefits were derived from the gathering of safety documents prior to the actual review of a manual. This procedure enabled the reviewer to incorporate related safety criteria into a manual along with the specific requirement. Methods of incorporation were a safety step or warning, caution or note placed in the body of the manual, and then to include the authority for the action (usually one of the documents in the compendium[1]) in the list of publications or referenced data table at the beginning of the manual.

Regarding warnings, cautions and notes this much should be said: MIL-M-21548 is the General Specification for FBM Weapon System Technical Manuals. It requires safety summaries at the beginning of the manual and states that warnings should emphasize an operation, procedure, practice, or condition that, if not strictly followed or maintained, could result in the injury or death of personnel. Cautions should be used when only damage to equipment is involved. For both warning and caution notices, a failure to comply with the instruction would result in some undesirable risk, the prevention of which is the reason for the notice. Notes should be (and were) used to emphasize important procedures or conditions. Notes, unlike warnings or cautions,

were located in the manuals prior to or following the work instruction. Warnings and cautions always preceded the work instructions.

Preliminary hazard analyses, as required by MIL-STD-882, System Safety Program For Systems and Associated Subsystems and Equipment, were prepared for the Poseidon missile only. Such an analytical procedure is a bonanza for the safety reviewer since in one document hazards are identified and an analysis already accomplished. Even for hazard analysis completed at the system level of detail, correlation of hazards with the step-by-step instructions in the equipment manuals was possible.

System Orientation Type Manuals

The purpose of Poseidon broad-scope manuals is to provide overall introduction and orientation. Information is presented in a general manner for familiarization. There were many of these manuals published and reviewed. Some had a brief description of the missile and subsystems. Others covered support facilities. Still others covered logistic concepts, missile configuration, and functional descriptions of missile prelaunch, flight sequences, maintenance, and system test. All types were reviewed for safety. It might well be asked what valid safety input could go into such manuals. The answer is that any program committed to safety must have requirements and concepts, not all of which fit into the format of working-type manuals. Understanding helps in complying with safety precautions and promotes accident-free operations. The system orientation type manual is an excellent vehicle for the safety specialist to develop safety awareness by requiring that safety rules and general principles be set down and explained. This was done on Poseidon.

Poseidon safety requirements reduce the possibility of ignition of propellant and explosive components during processing operations by a method of control known as the "Ignition System Missing Link." This requirement provides that key ignition or ordnance components are not installed at the same time in the system during processing. To assure compliance, two knowledgeable persons are required to be present during any movement or operation involving missing link components. The latter concept is known as the "rule-of-two," or "buddy system" for similar high risk operations. Both requirements were fully explained in the broad-scope manuals and can aid the reader in understanding the importance placed on safety. The reader was informed of unique explosive items in the system manuals, such as confined detonating fuse (CDF) and the safety confining caps that protect the worker. The Poseidon make-before-break grounding concept used on ordnance operations was explained at the orientation level. Safety steps would not be omitted at any processing level. Inadvertent activation, from electrostatic electricity, at the missile level would not occur. No opportunity was overlooked to include and/or explain the reasons for safety rules.

Ordnance Pamphlet (OP) 3666 is an example of the type of system manual under discussion. The title of this manual is Poseidon Missile UGM-73A, Missile System Analysis and Trouble Isolation, Submarines. The purpose of this OP is to describe the processing and maintenance of the missile and its related components on board the submarine. It does this in two volumes and nine parts. Volume 1 has a safety summary which was required by the safety reviewer on all Poseidon manuals. The safety summary is a list of every "warning" contained in the manual along with the statement that all personnel involved in the operation and maintenance of this equipment must fully understand the warnings and the procedures by which the hazard is to be reduced or eliminated. Motherhood? Possibly so. A "General Safety Precautions" paragraph was added to OP 3666, Chapter 1, that listed safety rules that must be observed by all personnel working on missiles and missile components on board submarines. Moreover, the introduction to these rules reads, "Injury or death can result from carelessness, failure to comply with approved procedures, or violations of WARNINGS, CAUTIONS, and safety regulations." The rules were written so they could be understood by submarine personnel (e.g., "Tools or other foreign objects must not be dropped between the missile and launcher tube. All objects must be removed from pockets, and tools must be secured to person or clothing before working in or above launcher tube").

System manuals were used to show a relationship between the Poseidon program and other comprehensive Navy safety programs in the Government documentation list used by the safety reviewer. Examples added to the reference table in OP 3666 were OP 4, Ammunition Afloat, and OP 3347, United States Navy Ordnance Precautions. Where the situation in OP 3666 called for broad safety precautions, the safety reviewer required a warning - "all applicable safety precautions shown in the table must be observed. Failure to comply may result in injury or death." The same technique was used on other manuals to tie in OP 3243, which is the Bureau of Naval Weapons General Safety and Industrial Hygiene program.

Poseidon manuals of broad scope were reviewed by senior safety experts with years of scientific and development background. Additionally, these experts were thoroughly familiar with the missile and had exposure to weapons system problems either first hand or from reviewing Trouble and Failure Reports sent back from the field. Only by such talent and experience can safety-significant details be emphasized in orientation manuals without giving the impression of pomposity or superabundance. There were still different schools of thought on how much and how often safety data should be included. Mutual concessions were made with the safety summaries. One school of thought believed summaries should elaborate on the warnings contained in the body of the manual and be placed in front of every hard binder or book part. Such elaboration would provide the safety rules or regulations for the book part being covered and repeat all rules, regulations, and warnings that were applicable to the entire manual. Those holding an opposing viewpoint did not see the need for summaries. A compromise was agreed on and it was decided that selected fleet manuals would contain safety rules and precautions in paragraphs, as mentioned previously, and certain shorebase manuals would contain safety rules and precautions at the start of work processing sections. All manuals would have a safety summary that listed the manuals' warnings.

Shorebase Equipment Manuals

For the purposes of this paper, equipment manuals are those detailed procedures that provide concise step-by-step instruction for preoperation, inspection, operation, troubleshooting and maintenance. They comprise the bulk of the manuals in Figure 2 and offered the biggest challenge to safety reviewing personnel. The award for complexity would have to go to OP 3667. This document, at last count, contains over 2197 pages of configuration, assembly, disassembly, processing, test requirements and procedures for the missile, special tools and surface support equipment. By the way, the 2197 pages did not include any classified pages or the guidance subsystem.

No single individual could have accomplished the depth of review required by safety on this manual, even if he had a working knowledge of all the principles covered in the compendium[1]. Completion of this reveiw required eight engineers for a three-week period. A comprehensive evaluation was assured by assigning to the safety reviewer those sections and chapters dealing with his primary field or experience. Review assignments were made directly from the OP index with no time lost in determining who should review what. Such a procedure was workable because missile subsystems and each component within each subsystem

were identified by the same reference designation generally used throughout the plant. Accomplishing the safety review for some of the individuals with limited shop and assembly exposure was not quite as simple as scheduling the assignments. Early in the review period, some safety reviewers found safety principles could not be imposed during one stage of missile processing without consideration of the subsequent operation. One safety rule that gave some trouble was the connecting of electro-explosive devices as the last operation in final assembly. Strict application of this principle may make the final assembly impossible or require a one-hand, blind hook-up by a midget, an operation that could be more of a potential hazard than a deviation from the general rule. Visual evidence in the form of a full scale mock-up of the Poseidon was used by the reviewers to achieve optimum safety. Where the subsequent operation, as verified on the mock-up, was impossible, or caused the worker to perform critical assembly operations without benefit of all his skills, the trade-off was made for the specific case over the general safety principle. However, the potential hazard was controlled in the subsequent operations with warnings and verification steps that assured that warnings were observed. The mock-up was useful for safety consideration of the maximum allowable weight to be handled by one man. Since the human factors standards of Ordnance Data (OD) 18413, Vol. 2, gives maximum allowable(s) in terms of height lifted from the ground, hardware feel and lift were used by the safety reviewer on border-line cases. In cases where the geometry of the hardware or particular location in the missile might strain a worker, the safety reviewer required a warning in OP 3667 that two men were necessary for the operation. Some slow-run-through processing of mock-up hardware disclosed probable misoperations that would not pose immediate personnel hazards but could jeopardize end product performance. This type of human error problem was flagged by caution notices such as, "Ensure that the flight control gyro package is oriented properly relative to the equipment section. Do not force the orientation pins into the wrong holes. Failure to comply could cause a flight malfunction".

The strict application by some technical writers and some safety reviewers of stereotype warnings in OP 3667 was not condoned by the safety reviewer. Specifically, volume usage of a warning notice similar to the following: "WARNING: Ensure that all system safety precautions and regulations are observed. Failure to comply may result in death." (in over 200 places in parts 2 and 3 of Vol. 3 and uncounted times in the other volumes) dilutes the effectiveness of the warning. The problem was solved by deleting it in any work section that was inherently a low risk of injury/death operation and, where a general-type warning was determined to be required, it was revised to specify where the precautions and regulations could be found. For example, WARNING: Comply with all applicable safety precautions contained in documents listed in table 1-1 to preclude injury. Of similar concern was the over emphasizing of warnings about industrial solvents. The blind use of certain warnings for trichloroethane, without any consideration for the operation, led some technical writers into a trap. To provide guidance where it was needed, the safety reviewer recommended a change in the warning notice, "Do not use trichloroethane in any unventilated areas without wearing a respirator. Prolonged inhalation may result in death. Wear protective gloves and clothing as contact with fluid will cause skin injury." This notice is of questionable value for this industrial solvent used in open areas for damp cloth wiping of small surfaces. Trichloroethane is the preferred solvent for hand solvent wiping or brushing, because it is relatively non-flammable and has low toxicity. However, if there is to be prolonged exposure to the trichloroethane vapor in an unventilated area, the personnel protection should be a self-contained breathing apparatus or air supplied by means of a pump and hose. Since profuse solvent usage in an unventilated area was not likely for the damp cloth wiping operation ashore, the warning was changed to: "Do not use trichloroethane in an unventilated area. Do not inhale vapor. Prolonged inhalation of vapor may result in death." Substitution of the latter warning for the former warning was the recommendation of the safety specialist.

Safety review comments included complaints about errors of commission as well as omission in basic processing procedures covering missile checkout. Where information in the manual was missing, the safety recommendation usually went as follows: "Page 11-1, paragraph 11-2-2; Add general information about all emergency shutdown procedures such as: How they are invoked (i.e., System Test Emergency Shutdown Procedure is required in the processing work segment): How or if emergency shutdown procedures apply to processor, semi-automatic and operator mode selection. The alarms/indicators that require an emergency shutdown procedure should be identified."

Shorebase procedures that dealt with Fleet-return missiles were scrutinized for control of potential hazards due to critical items inadvertently left on the missile. The safety reviewer required an inspection to assure that all batteries were removed and a warning - "an activated battery is a hazard, wear protective clothing and equipment. Battery electrolyte can cause severe burns and blindness." Ordnance components were controlled most diligently. Each live component was identified and inspection required to assure its removal. Inspection instructions for the "missing link" hardware were supplemented with notes to increase safety awareness and provide the worker on the floor with knowledge of unique Poseidon requirements. A note for Inverters read, "Inverters are missile link components. The rule of two shall be invoked during any handling operation." Reference tables were used to show the document which implemented safety requirements and were tied into the ship-shore interface. The in-depth review of equipment-type manuals by the safety reviewer disclosed some problems regarding the level to which the documents were written. For the most part, these problems were minimal because the customer's requirements were for MIL-M-21548, and this specification states that technical information shall be written for comprehension by an enlisted technician who has had some formal Navy training in the applicable technical field. Safety instructions and warnings to use "make-before-break" and "one hand hook-up rule" were acceptable to all organizations. Other instructions in the manual, although understandably complex, were not approved by safety if they could be misused. For example, a maintenance instruction covering metal surface refinishing usually went - "Remove scales, oxides and flaked paint with scraper or wire brush. . ." This, in the judgment of the safety reviewer, gave the worker an option to use a "wire brush" without stating whether it were a hand brush or a power wire brush. In the latter instance, safety would require that eye protection be worn by the operator to prevent eye injury from flying particles. The problem was solved by identifying the tool as a hand wire brush, where the possibility existed that the worker would use the power tool. When the experience of the safety reviewer was limited, he contacted his counterpart in the field for information about the available hand tools.

Safety review of equipment manuals can provide, or at least supplement, a figure of merit as to how comprehensive the hardware design reviews were. A case in point is the Surface Support Equipment Covers and Mats Manuals (OD 42595). Sharp corners on a CDF protective cover posed a negligible personnel hazard when used on the vertical missile. Use on a horizontal missile was not ruled out in the manual, and such use presented a potential puncture, cut, or scrape hazard. The safety reviewer required that a warning notice be inserted in the manual to advise the worker to avoid the sharp points. The identification of this potential hazard was sufficient reason for a subsequent design change that eliminated the hazard.

System design reviews can be subjected to the same figure of merit, provided the safety review is accomplished in accordance with the same guidelines used on the Poseidon program. To illustrate this point, the Thrust Vector Control (TVC) Plumbing Assembly Manual OD 43123 is worthy of note. The TVC plumbing assembly is used to control the flow of pressurized hydraulic fluid to the first and second stage TVC during the Poseidon motor checkout. This plumbing assembly is the last equipment in a three equipment system, and it does not contain any relief device. Equipment number one is a pressurization unit, and equipment number two is a filtration stand in which there is a relief valve for the system. Interconnection of equipments is made by flexible hoses. The safety reviewer determined for a worst case analysis a maximum pressure allowed for the flex-hoses could be exceeded. An analogous situation would exist if the pressurization unit (first equipment) were other than the design engineer envisioned, i.e., a substitute unit. Safety recommendations considered both possibilities and resulted in flex hoses having higher ratings along with a warning notice in the OD as follows: "WARNING: Do not connect plumbing assembly to any external pressure source not regulated to 2,900 psig. Exceeding 3,500 psig maximum operating pressure rating of plumbing assembly hoses could cause injury."

The value of the list of safety documents applicable to Navy Systems and Ordnance should be stressed as an aid in the review of equipment-type manuals. It is questionable that even an experienced engineer would have the presence of mind to require a copy of the drivers' handbook, 3 reflector flares, copies of the Depot drivers regulation and an operator's report of motor vehicle accident (Form 91) in the equipment manual for the missile straddle carrier. However, an engineer with sufficient foresight to read the requirements of OP 2239 (which is the Drivers' Handbook compiled specifically for drivers of Navy vehicles engaged in the transportation of ammunition, explosives and other dangerous articles (AEDA) intra-station and over public highways) would have found such safety requirements listed. The safety reviewers on Poseidon manuals had this kind of foresight and were able to review the related safety requirement. Safety items for the straddle carrier were included in the review results. Other benefits were derived from selectively reviewing government safety documentation prior to the review of a technical manual. This procedure provided our customer with a means of tying his numerous safety programs into Poseidon. Such integration was accomplished by providing for the specific safety requirement in the body of the manual and citing the authority for the requirement in the table of publication. By faithfully following such a procedure during the safety review, it can be certified that no significant authoritative safety document was precluded from Poseidon Technical Manuals.

Fleet Technical Manuals

Rules, regulations, and the documents in which they can be implemented change from the plant, to the base, to the fleet. Some equipment and some manuals do not. One of the problems detected by the safety review was that a potential hazard previously controlled by a shorebase requirement was not necessarily controlled by the same publication at sea. Compounding the problem early in the program was the lack of a cross index and there was a small annoyance factor of having another contractor responsible for the procedural manuals. It must be said in all candor that the safety review of Poseidon technical manuals was not the universal formula to solve all problems. A statement of fact is that the safety review contributed significantly to the detection of such problems and that safety personnel assisted the customer and publication personnel in furnishing the most prudent solutions.

The Poseidon missile underwent a series of shore-based launches before any tests were conducted at sea. Such a schedule required a review of the shore-based operational manuals ahead of the fleet manuals. Potential hazards already identified in the shorebased review were usually checked first by the fleet manual safety reviewer with some interesting results. For example, the potential hazard from an activated missile battery was controlled on shore by the use of warnings and emergency procedures during the packaging and unpackaging operation. The procedure read as follows: "Warning: Wear approved eye protection or face shielding and protective clothing when handling activated battery. Activated battery generates heat, releases toxic gas, and could liberate chemicals. Contact of battery chemicals with skin may cause burns. Clear area of all personnel, except those wearing protective clothing and safety equipment." The same emergency procedure was considered for OP 3751, which is the missile standard maintenance procedure aboard the SSBN. However, the safety reviewer found he could not invoke the shore-base manual as the implementing document, because shore-base manuals are not applicable afloat. Rationale for such treatment became apparent to the reviewer as the customer's safety requirements to protect the SSBN environment became known. Furthermore, new potential hazards were added by the SSBN requirements for like operations on shore. Specifically, there were problems with solvents. Methyl chloroform (trichloroethane) was used to clean "black boxes" ashore. The operation was preceded by a warning notice regarding ventilation and for such an operation the hazard was controlled. This same operation on board the SSBN required isopropyl alcohol which added a potential fire hazard that had to be considered by safety review and resulted in a warning: "Keep isopropyl alcohol away from sparks, heat or flame, keep container closed and avoid prolonged breathing of vapor."

At this point in our case history, some information about the method of documenting the safety recommendations on Interdepartmental Communication (IDCs) and the full scale hardware mock-up should be recorded. The principal IDC reason was to preserve a historical record of the safety reviewer's findings. Towards the end of the task, retention of the manual reviewed, along with a copy of the IDC, was found to be desirable. Because of extensive changes to some manuals, it was difficult to find the operation for which the safety comment was made. Keeping the manual was easier than including the before and after operation in the IDC. On the positive side, a carbon copy of the IDC was excellent advertisement of the quality of the systems safety groups' work. A carbon copy also proved to be an effective tool to get action started by line organizations when their special interest might be affected.

Some potential personnel hazards would not have been detected so readily, were it not for the mock-up of the SSBN launch tube and missile. The capability of working through the procedures in the manual at hand and, under similar conditions, in the field was an invaluable aid to the safety reviewer. Where the SSBN manual required protective covers over missile igniters during certain operations, a visit to the mock-up was in order. The results of the visit showed the covers were subject to man handling because of the tight quarters. Interface possibilities such as these, which could lead to potential personnel hazards were controlled in the manual with precautions to install the igniter cover with great care in order to avoid impacting or scraping ordnance. The visit to the mock-up also showed the safety reviewer that the heavy launch tube door could be a potential hazard due to motion of the sea. A warning was added to the manual requiring the tube door to be lock-pinned open to prevent injury should it swing. In summation, the full scale mock-up illustrated ship-shore interface problems and provided the safety reviewer with tangible evidence of the validity of his conviction. He could impose safety controls with confidence.

Conclusions

A case history-type paper does not always provide a flow of facts from which conclusions are evident. Such is the nature and intent of this report. However, certain

observations can be listed with the completion of the Poseidon technical manuals' safety review. Nothing in this paper invalidates the statements below.

1. A literature search resulting in a compendium of all safety rules and regulations is the essential first step for any product mix like the Poseidon FBM system.
2. Senior people with proven achievements in the safety discipline must direct the effort for a minimum hazard risk without superabundance of safety controls.
3. Hardware or full scale mock-up must be available to affirm or refute safety problem at the worker and hardware interface.

Reference

1. Compendium: This is a bibliography compiled by the author, that identifies safety documents with title and a brief comment. It is much too long for this type paper. However, a copy will be provided on request to the author, at Lockheed Missiles & Space Company, Inc., MSD, Dept. 84-13, Bldg. 182, P.O. Box 504, Sunnyvale, California 94088.

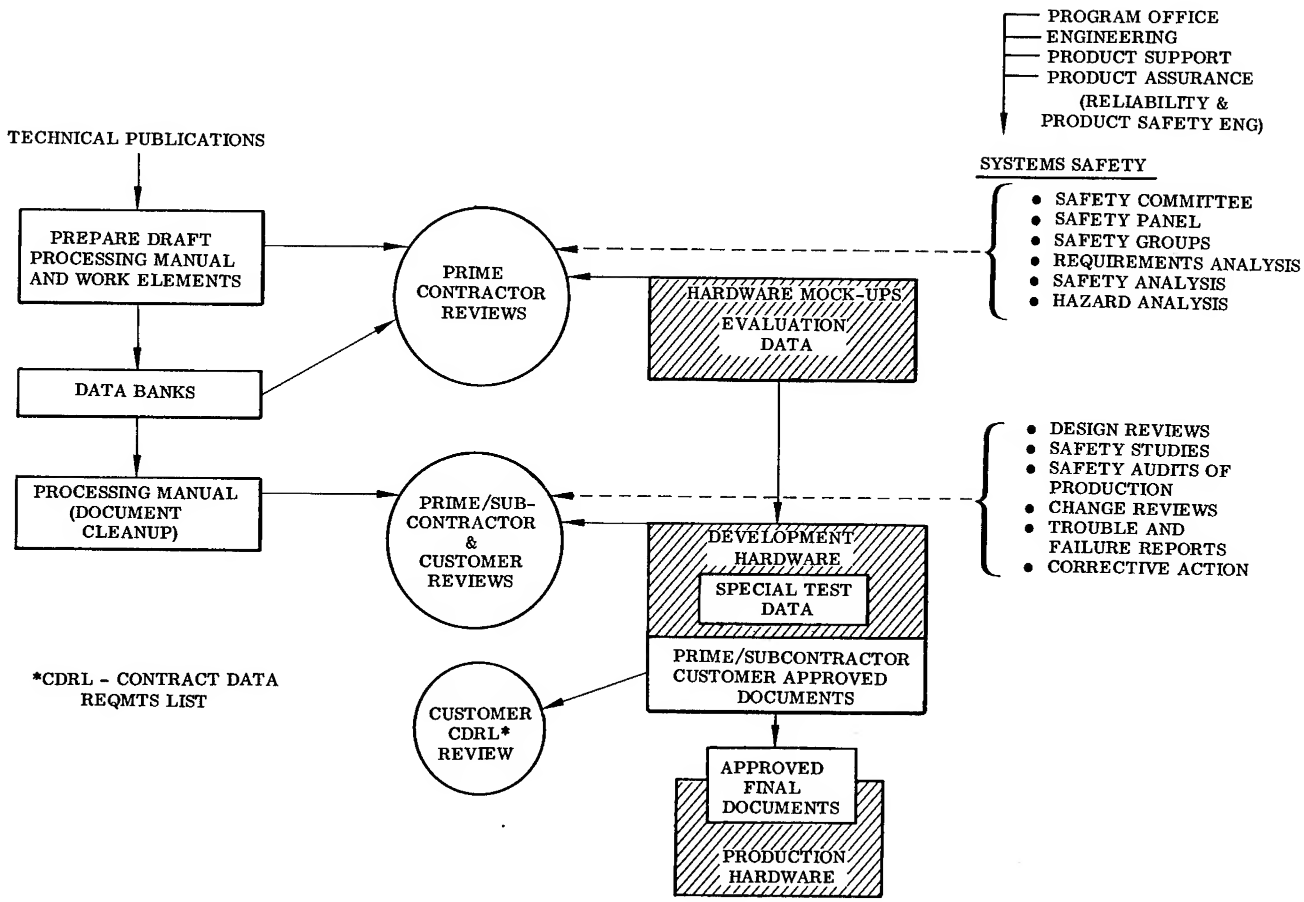


Figure 1 -- ORGANIZATION, MAJOR FUNCTIONS AND INTERFACES

FLIGHT TEST MANUALS

LMSC-D052060

- POSEIDON C3P-PEM FLIGHT TEST VEHICLE DESCRIPTION
- VOLUME 1 BOOSTER, DELIVERY, AND INSTRUMENTATION SYSTEMS
- VOLUME 2 DEVELOPMENT RE-ENTRY SYSTEMS
- SUPPLEMENT PARTS 1 AND 2 POSEIDON C3P-PEM FLIGHT TEST VEHICLE DESCRIPTION

LMSC-808640

- POSEIDON C3X FLIGHT TEST VEHICLE DESCRIPTION
- VOLUME 1 BOOSTER, DELIVERY, AND INSTRUMENTATION SYSTEMS
- VOLUME 2 DEVELOPMENT RE-ENTRY SYSTEMS
- SUPPLEMENT PARTS 1 AND 2 POSEIDON C3X FLIGHT TEST VEHICLE DESCRIPTION

MISSILE DESCRIPTION MANUALS

OP 3662

- POSEIDON MISSILE UGM-73A MISSILE SYSTEM ORIENTATION

OP 3663

- POSEIDON MISSILE UGM-73A MISSILE SYSTEM DESCRIPTION
- VOLUME 1 BOOSTER AND DELIVERY SYSTEMS
- VOLUME 2 EVALUATION AND OPERATIONAL TEST SYSTEMS
- ADDENDUM A POSEIDON MISSILE SCHEMATIC DIAGRAM
- ADDENDUM B POSEIDON MISSILE FUNCTIONAL BLOCK DIAGRAM

SPECIAL DATA

OP 3751

- STANDARD MAINTENANCE PROCEDURES, SSBN

OP 3689

- STANDARD MAINTENANCE PROCEDURES, TENDER

OD 42810

- POSEIDON MISSILE UGM-73A PROCESSING DOCUMENTATION PLAN, SHOREBASE

OD 42910

- POSEIDON MISSILE UGM-73A POMFLANT MANUAL VERIFICATION PLAN

OD 43242

- DOCUMENTATION REQUIREMENTS FOR OPERATION AND FLEET MAINTENANCE OF SSBN 627 CLASS AFTER OVERHAUL

- MAINTENANCE BULLETINS
- SPALT DOCUMENTS
- PREVENTIVE MAINTENANCE MANAGEMENT PROGRAM DATA
- ATOMIC WEAPONS RETROFIT ORDERS (AWROs)

MISSILE SYSTEM ANALYSIS & TROUBLE ISOLATION MANUALS

OP 3664

- POSEIDON MISSILE UGM-73A, MISSILE SYSTEM ANALYSIS AND TROUBLE ISOLATION, SHOREBASE
- VOLUME 1 MISSILE SYSTEM TESTS
- VOLUME 2 MISSILE PACKAGE TESTS

OP 3665

- POSEIDON MISSILE UGM-73A, MISSILE SYSTEM ANALYSIS AND TROUBLE ISOLATION, TENDERS
- VOLUME 1 MAINTENANCE CRITERIA AND CORRECTIVE ACTION
- VOLUME 2 MISSILE SYSTEM TESTS
- VOLUME 3 MISSILE PACKAGE TESTS

OP 3666

- POSEIDON MISSILE UGM-73A, MISSILE SYSTEM ANALYSIS AND TROUBLE ISOLATION, SSBN
- VOLUME 1 MAINTENANCE CRITERIA AND CORRECTIVE ACTION
- VOLUME 2 MISSILE SYSTEM TESTS

MISSILE PROCESSING MANUALS

SHOREBASE

OP 3667

- POSEIDON MISSILE UGM-73A, MISSILE PROCESSING, SHOREBASE
- VOLUME 1 PROCESSING REQUIREMENTS SUMMARY
- VOLUME 2 PROCESSING SUPPLEMENTARY DATA
- VOLUME 3 PROCESSING WORK SECTIONS

TENDER

OD 43145

- STANDARD OPERATING PROCEDURES, AS-33 CLASS TENDER

OD 43146

- STANDARD OPERATING PROCEDURES, OPERATIONAL TEST MISSILES

SSBN

OD 43144

- STANDARD OPERATING PROCEDURES, FBM WEAPON SYSTEM SSBN-627-CLASS

RE-ENTRY VEHICLE PROCESSING, HANDLING & TESTING MANUALS

SWOP W68.76-0

- WEAPONS SUMMARY, W68-0/MK 3 RE-ENTRY BODY ASSEMBLY

SWOP W68.76-1

- ASSEMBLY, TEST, STORAGE AND MAINTENANCE PROCEDURES, W68-0/MK 3 RE-ENTRY BODY ASSEMBLY

SWOP H68

- OPERATION AND MAINTENANCE INSTRUCTIONS, HANDLING EQUIPMENT, W68-0/MK 3 RE-ENTRY BODY ASSEMBLY

SWOP H68A

- ILLUSTRATED PARTS BREAKDOWN, HANDLING EQUIPMENT, W68-0/MK 3 RE-ENTRY BODY ASSEMBLY

SWOP T68

- OPERATION AND MAINTENANCE INSTRUCTIONS WITH ILLUSTRATED PARTS BREAKDOWN, TEST EQUIPMENT, W68-0/MK 3 RE-ENTRY BODY ASSEMBLY

MISSILE TEST AND READINESS EQUIPMENT MANUALS

OP 3670

- MISSILE TEST AND READINESS EQUIPMENT MK 6 MOD 1, DESCRIPTION, OPERATION AND MAINTENANCE

OP 3671

- MISSILE TEST AND READINESS EQUIPMENT MK 7 MOD 2, DESCRIPTION, OPERATION AND MAINTENANCE
- VOLUME 1 DESCRIPTION
- VOLUME 2 OPERATION AND MAINTENANCE
- VOLUME 3 SYSTEM VERIFICATION

OP 3672

- MISSILE TEST AND READINESS EQUIPMENT MK 7 MOD 3 DESCRIPTION, OPERATION AND MAINTENANCE
- VOLUME 1 DESCRIPTION
- VOLUME 2 OPERATION AND MAINTENANCE
- VOLUME 3 SYSTEM VERIFICATION

OD 42429

- MISSILE TEST AND READINESS EQUIPMENT, MK 6 MOD 3, SYSTEM ACCEPTANCE TESTER, DESCRIPTION, OPERATION AND CALIBRATION

OD 42430

- MISSILE TEST AND READINESS EQUIPMENT, MK 6 TYPE III MODULES ACCEPTANCE TESTER, DESCRIPTION, OPERATION AND CALIBRATION

OD 42431

- MISSILE TEST AND READINESS EQUIPMENT, MK 7 MOD 2 SYSTEM TESTER, DESCRIPTION, OPERATION AND MAINTENANCE INSTRUCTIONS

OD 42432

- MISSILE TEST AND READINESS EQUIPMENT, MK 7 MOD 3 SYSTEM TESTER, DESCRIPTION, OPERATION AND MAINTENANCE INSTRUCTIONS

MISSILE SURFACE SUPPORT EQUIPMENT MANUALS

OD 42893

- POSEIDON MISSILE UGM-73A SYSTEM SUPPORT EQUIPMENT

Figure 2 -- POSEIDON TECHNICAL MANUALS

APPLICATION OF SAFETY DISCIPLINES TO SRAM PROGRAM

Ronald D. Craig
SRAM System Safety
The Boeing Company
Seattle, Washington

I. Introduction

This paper will discuss the successful application of system safety disciplines to a missile (SRAM) program. SRAM (Short Range Attack Missile) has been designed, developed and evaluated by The Boeing Company for the Air Force's Aeronautical Systems Division at Wright-Patterson Air Force Base, Dayton, Ohio. The nuclear missile is a strategic weapon planned for use on the FB-111 fighter-bomber and late model B-52's. It is designed to be launched from these strategic Air Force bombers against ground targets. Since it is a rocket-propelled air-launched missile that can fly at supersonic speeds, it provides a stand-off capability which will assist in the penetration of sophisticated enemy defense systems.

Safety Management—Application of safety disciplines to the SRAM program first required the establishment of a safety management philosophy to ensure that hazards were identified and that actions were initiated to prevent or control identified hazards. The management philosophy successfully used on the SRAM Safety Program was to establish System Safety as an active element in the engineering organization providing a continuous System Safety examination of design, test planning, testing, production and operational phases of the program. It is important to note that the SRAM System Safety organization was established as an active member of the overall SRAM engineering team with specific program objectives and design requirements to meet. Qualified personnel were selected to penetrate all phases of the system design and development.

After more than five years of following this management philosophy SRAM has completed a design/development phase, including a successful flight test phase, fabrication of production hardware, and successful activation of SRAM strategic Air Force bases, which in total represent in excess of 6 million manhours without experiencing *any* catastrophic or critical accident or incidents to personnel or equipment.

The other basic approach to safety management is one in which the safety organization is in a quasi staff position, where the major portion of safety activities are performed in piece meal fashion by different project organizations. This type of a management approach has the disadvantage of not providing the safety organization with the continuity of review and the depth of knowledge necessary to identify and correct safety hazards.

The point that is important here is again the team concept. The successful team always has each member doing his assigned task to accomplish the total team (project) objective, and system safety on the SRAM Program was an active part of the team. Just as a football team cannot succeed with all quarterbacks, a project cannot succeed with all design engineers.

Plan—The next step in the SRAM program was to develop a System Safety Program Plan for both the Design, Development, Test and Evaluation (DDT&E) and production programs which outline tasks, methods, and responsibilities to meet the program objectives and requirements. These plans were then coordinated and approved by Aeronautical Systems Division (ASD) at Wright-Patterson Air Force Base. After approval these plans were executed by the safety organization.

Requirements—In general the SRAM program safety requirements for the DDT&E phase were to eliminate and/or control all category III and IV hazards to a level not to exceed 1.2×10^{-4} per missile launch. In addition there is a nuclear safety requirement to meet and design the system to meet the requirements of Nuclear Systems Safety Design Manual (AFSCM 122-1), and that unexpected events involving nuclear weapons shall not contribute more than 1×10^{-8} to the total critical and catastrophic hazards. The DDT&E safety program was conducted in accordance with MIL-S-38130, which was basically a safety analysis identification standard, without specific direction or guidance for Production and Operational programs. The SRAM Production safety program is designed to meet the requirements of MIL-STD-882, which basically extends safety into the production/operational phase of a program. The basic difference in shifting from a DDT&E program controlled by an analysis oriented standard into a production program controlled by a different standard extended to include production, is one of shifting from a hardware analysis to a personnel/hardware/procedure interface task oriented analysis. The production phase analysis methods are more tailored to people/procedure problems, in relationship to hardware failure modes that are primarily addressed in the design phase.

It is important to note that the continuity of first designing safety into the product, followed by assurance that designed safety features are not compromised by the testing program, and that unsafe operational procedures are not used must be examined continuously.

The sections to follow in this paper will show the analytical techniques applied to both the DDT&E and Production programs, how they differ, and the reasons for the difference. The paper will conclude with a discussion on manning and relative cost of the safety program.

II. Analytical Techniques Applied to Design, Development and Testing Program Phase

The safety analyses conducted on the DDT&E program, which in general encompassed four and one-half years followed the guidance provided by MIL-S-38130. The SRAM program was one of the first programs to have a numerical safety requirement specified as a firm design requirement for hardware development. The establishment of this numerical requirement established the need to develop an analytical technique beyond that specified in MIL-S-38130. This analysis technique is a computer math model fault tree simulation program designed to show that the SRAM System Safety numerical specification number was met. Figure 1 shows the relationship of the various safety analyses performed and Figure 2 shows their relative phase relationships.

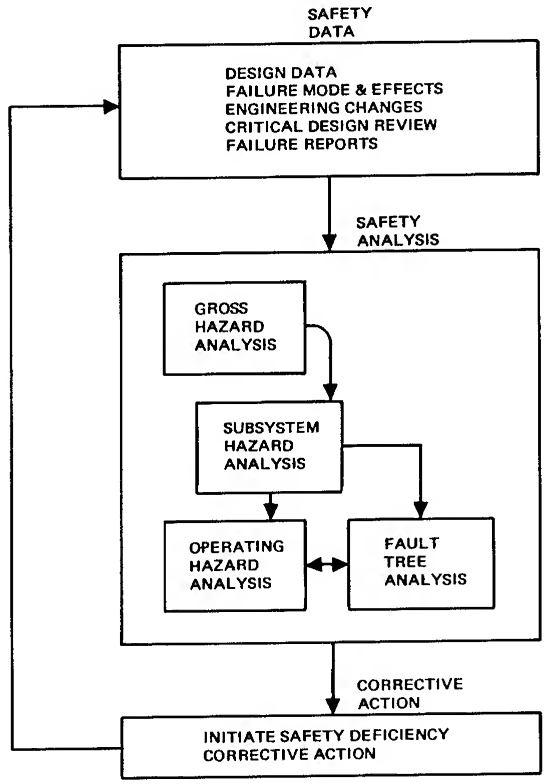


FIGURE 1

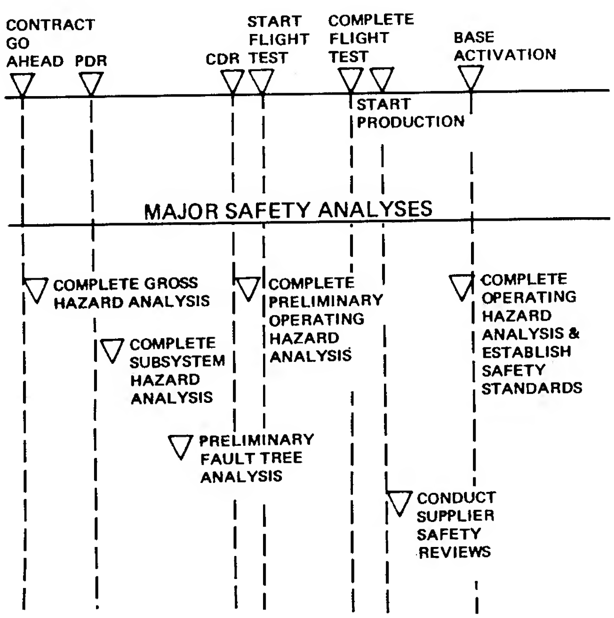


FIGURE 2. MAJOR PROGRAM MILESTONES

Gross Hazard Analysis—This analysis was completed early in the DDT&E program to identify system "undesired events" which are classified as conditions which are Catastrophic and/or Critical. A management decision was made at this time in conjunction with ASD, that both catastrophic (IV) and critical (III) hazard conditions would be lumped together and the total would not exceed 1.2×10^{-4} per missile launch. Realistically, it is not practical to attempt to distinguish between catastrophic and critical events in an analytical process. The gross hazard analysis resulted in 8 "undesired events." Identification of these are shown in Figure 3.

- PREMATURE MOTOR IGNITION
- PREMATURE W/H ARMING & FUZING
- MOTOR OVERPRESSURE
- MISSION ABORT HAZARDS
- MISSILE FIRE
- PREMATURE MISSILE RELEASE
- MISSILE/CARRIER COLLISION
- PERSONNEL INJURY

CURRENT SAFETY ASSESSMENT

REQUIREMENT	ASSESSMENT
NON-NUCLEAR 1.2×10^{-4}	8.2×10^{-5}
NUCLEAR 1×10^{-8}	3.7×10^{-10}

FIGURE 3. GROSS HAZARD ANALYSIS CLASS III & IV HAZARDS

Subsystem Hazard Analysis—This analysis was performed almost in parallel with the gross hazard analysis, which in retrospect was an error. After examination of the results of this analysis it was concluded that a delay in the start of this analysis would have been more cost effective to the program. In general the subsystem hazard analysis is a failure mode analysis on hardware subsystems. Because a large number of the subsystems for the SRAM system are supplied by subcontractor, the subcontractors were directed to perform the analysis. However, without specific direction from the SRAM safety organization the subcontractor performed a detailed analysis on all failure modes. It would have been more cost effective if the gross hazard analysis would have been completed, then extend that analysis to specify specific "undesired events" for each subcontractor, and then have analysis performed only on specific hardware that could effect those selected events. It was also our experience that many of the subcontractors did not have safety engineers in their organization, which in turn required more coordination on Boeing's part to obtain an effective analysis. In the future it would be advisable to plan the analysis required by the subcontractors more effectively, and provide better instruction for performing subsystem hazard analysis.

Fault Tree Analysis—This analysis technique has been discussed many times in various papers and this author will not discuss the technique, other than to point out that the SRAM program used the technique as an effective tool early in the program to accomplish some specific objectives, but the fault tree was not looked upon as an end in itself. The SRAM Safety organization did develop a math model simulation program to calculate the probability of the occurrence of rare events.

A unique feature of the SRAM fault tree analysis method was the use of a phase independent fault tree. Because of the numerical requirement to prove the SRAM system meets specific requirements during various phases in a typical mission, it was necessary to develop a phase oriented numerical analysis technique. Figure 4 shows a typical SRAM mission profile which represents a Stockpile-Target-Sequence (STS). The SRAM fault tree was constructed independent of the phases shown in Figure 4 so that each phase could be simulated independent and combined to represent a missile launch, as required by the system specification. Figure 5 shows an example of a fault tree segment which has been drawn as a phase independent tree. In order to simulate the actual flight of the missile the safety analyst provided inputs into the computer program to simulate actual equipment operation. Figure 6 shows the computer input format. The fault tree in Figure 5 is analyzed by the safety engineer. For example, the relay contacts have a failure rate (λ), with a probability of $P \simeq \lambda Kt$. The K factor accounts for environmental considerations. The remaining parameter time (t) represents the time duration of the phase. The codes used provide the computer the necessary information to calculate the probability of each event occurring for each phase. The computer program with the use of importance sampling determines the probability of the hazard condition occurring

during each phase. The computer program will print out the critical path of each phase showing the critical components. At the completion of each run a program summary is provided which gives the total probability of occurrence for the missile launch (mission) together with the probability for each phase and the critical paths. The fault tree analysis method provided an important baseline from which many design changes were made to improve the safety of the SRAM system. It is very important to note that the numerical requirement provided the safety organization an effective tool to force design changes based upon probability data to show safety improvement to the system.

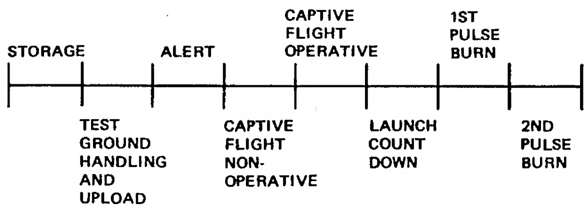


FIGURE 4. SRAM STOCKPILE-TARGET-SEQUENCE

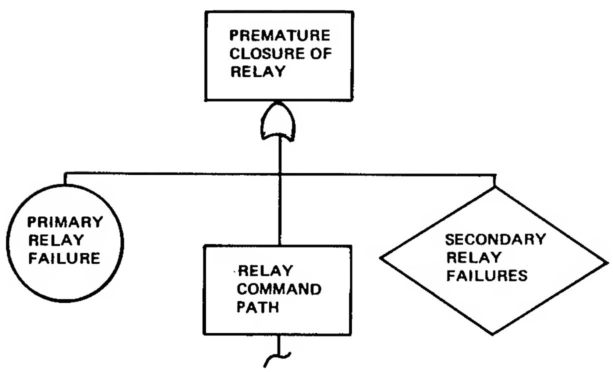


FIGURE 5.

Operating Hazard Analysis—A preliminary Operating Hazard Analysis was performed during the early DDT&E phase. This analysis was conducted primarily on hardware, prior to test procedures and Technical Orders (T.O.) being written. The purpose was to examine hardware in its operational mode and provide potential hazards to the test procedure and T.O. personnel such the caution and warning notes could be added to the test procedures.

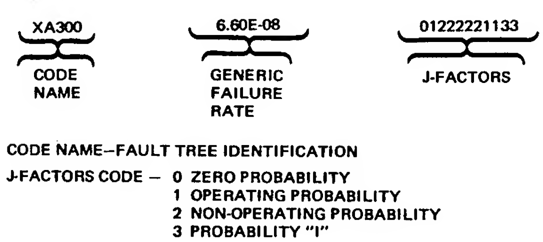


FIGURE 6

III. SRAM Flight Safety Program

Early in the planning phase for the flight test program it became apparent that the successful continuation and deployment of the SRAM system was very dependent upon a successful flight test demonstration. It was also apparent that a major accident involving the SRAM/carrier would seriously jeopardize the SRAM program. Therefore, a flight safety program was set up to assure the safe completion of the flight test program. The outline of this program is shown in Figure 7.

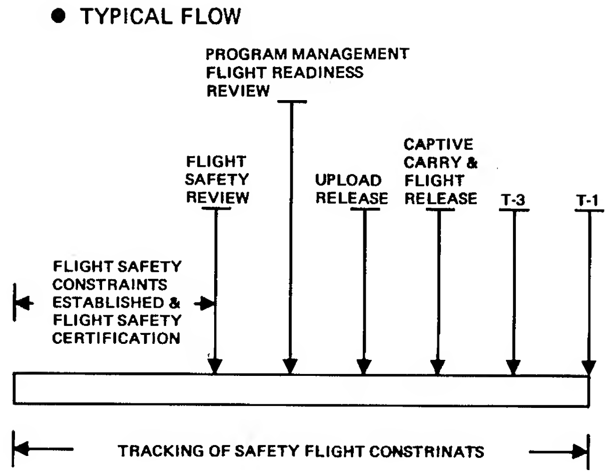


FIGURE 7. FLIGHT SAFETY PROGRAM

Flight Safety Certification—The first step in conducting the flight safety program as shown in Figure 7 was flight safety evaluation of SRAM equipment installed or operated in carrier aircraft during the test program. Evaluation is based on data from qualification tests i.e., vibration, shock, EMI, proofload, ordnance evaluation, and explosive atmosphere, essential to demonstrate flight safety. Based upon successful completion of these tests Flight Safety Certification is provided for each flight.

Flight Safety Constraints—were established for each flight. These constraints were established against upload, captive carry, and launch of the missile. The constraints were reviewed at the flight safety review, and the program management flight readiness review. After all constraints for upload, captive carry and flight release had been removed System Safety provided a release letter to allow for use of the system.

Safety Reviews—were conducted as indicated in Figure 7. In the safety review each flight was examined for specific flight hazards associated with the selected carrier/missile flight profile. The probability of Critical/Catastrophic events occurring during the flight was calculated and evaluations were made on the risk of the mission. In addition many special studies were requested by the flight crew and the results of these studies were examined during the safety review. Examples of these studies will be discussed later in this paper. For the first several flights there was a separate flight safety review held prior to the management review. However, as the flight test program became more mature, only the management review was held, which included a safety review item on the agenda.

Flight Safety Working Group—became active after the completion of the program management review. This group provided an overview of the remaining safety constraints at a high management level to assure the timely and safe resolution of all constraints. The T-3 and T-1 reviews were final reviews to assure that all equipment/procedures/personnel were ready for the scheduled flight. There were 38 missile launches during the flight test program which were reviewed by this method, without a single accident/incident that caused injury to personnel or major damage to equipment.

IV. Safety Review of Engineering Changes and Testing Program

During the course of both the DDT&E and Production programs the safety organization actively reviewed all engineering changes for possible impact on System Safety. For

those changes which had safety impact trade studies were conducted to examine alternatives, and in some cases changes were not approved because of an unacceptable impact on safety. In these evaluations the use of the fault tree numerical assessment was invaluable to show the effects of various approaches and to reach a final decision on the impact on the required numerical assessment. In addition to changes the safety organization reviewed test plans, and test procedures to assure that planned testing would not compromise designed safety features, and that test personnel were not placed in a hazardous condition.

The safety engineer also participated in the evaluation of equipment failures, both at the system, subsystem, and component level. The safety engineers were an active part of critical parts evaluation in terms of hardware "physics of failure" investigations. Safety engineers attended design reviews and examined and commented, and changed component layout, wire routine, pin assignments, and application and type of insulation material.

V. Safety Corrective Action

The most important element in any safety program is the correction of safety deficiencies. Safety is not achieved by the mere performance of a safety analysis alone. In the evolution of the System Safety discipline the development of analytical techniques have far exceeded the development of methods to correct safety deficiencies. Many organizations and engineers have become fascinated with development and modification of analysis techniques without regard to how they can be applied to a program to obtain the correction of safety problems.

The SRAM System Safety organization as a part of the overall design team was able to initiate and have incorporated a large number of hardware/software/procedure changes throughout the conduct of the DDT&E and Production program. Examples of these changes are: modification of the motor arming and ignition circuits, changes to launch countdown sequence to allow for additional testing of flight control system prior to release, addition of interlock circuits for booster testing to prevent application of power to booster if motor is connected, safety interlocks on missile rotary launcher to protect test/maintenance personnel; and many more. These corrective actions were accomplished due to the depth of knowledge of the safety engineer who with the use of the fault tree numerical analysis method was able to show potential problem and the impact on the numerical safety requirement.

Much more work has to be done in this area to improve the method by which corrective action is initiated together with follow-up action. Several methods have been considered by the SRAM organization for tracking safety problems, however, any type of tracking system requires additional people and budget and corrective action planning should be included and funded early in the program development phase.

VI. Special Safety Studies

During the course of the SRAM program the system organization received a number of requests to conduct special studies. These studies were supplemental to the analysis and were conducted to answer special questions under a specific condition. There were several studies completed to examine and predict probability of missile/carrier collision. These studies examined the missile and carrier for selected launch conditions and analyzed carrier/missile faults which could result in collision at missile launch or collision of missile/carrier in free flight. This paper will discuss two examples of special studies that were conducted as a result of requests by the flight crew. Early in the flight test program concern was expressed over possible damage to the missile carrier aircraft due to motor deflagration upon ground impact and resultant motor fragment damage to the aircraft. There was also concern over missile break-up and possible ricochet of parts which might collide with the aircraft. The following examples will show the methods and results of these studies.

Motor Deflagration (Explosion)–Analysis was conducted to determine under what conditions the motor would deflagrate. These analyses took the form of critical diameter analysis which determine the sensitivity of the motor to deflagrate. These analyses resulted in showing possible impact velocities which could result in motor propellant reaction. In general a missile launched or jettisoned from an aircraft could under worst case conditions result in a motor deflagration upon ground impact. The resulting deflagration will have a fragment pattern which could be a hazard to the aircraft under certain flight conditions. The degree and nature of this hazard was to be determined and reported upon prior to the flight of missile. Figure 8 shows a summary of fragment density and impact velocity studies which were conducted to determine the type of fragments and velocities from a deflagrating motor. From these data penetration studies of the aircraft and maximum height of the fragments were analyzed for possible aircraft engine ingestion or damage. The

WT (LB)	DRAG COEF * K (FT^{-1})	DENT VELOCITY (DAMAGE CRITERIA) (FT/SEC)	DENT MIN SEPARATION DIST (FT)	DENT MIN SEPARATION TIME (SEC)
1**	0.0158	785	26	0.03
1	0.0158	785	NO DENT	
4	0.01	500	20	0.05
8	0.008	392	55	0.12
15	0.00645	318	100	0.24
20	0.00585	288	120	0.30

PENETRATION VELOCITY (HAZARD CRITERIA) (FT/SEC)	PENETRATION MIN SEPARATION DIST (FT) DIST (FT)	PENETRATION MIN SEPARATION TIME (SEC)	MAX FRAG HEIGHT (FT)	MAX FRAG HEIGHT TIME (SEC)
2,400	NO PENETRATION		181	1.75
2,400	NO PENETRATION		150	1.6
945	NO PENETRATION		215	2.18
600	NO PENETRATION		255	2.4
393	70	0.15	301	2.76
324	104	0.24	324	2.8

* FLAT PLATE
** 1,000 FT/SEC (INITIAL VELOCITY)

FIGURE 8. FRAGMENT DENSITY AND IMPACT VELOCITY

table in Figure 8 shows the minimum distance for penetration and/or denting the aircraft. Figure 9 shows the maximum height of the motor fragments. From this data specific aircraft flight profiles could be examined to determine if any hazard existed to the launch aircraft, photo plane, or chase aircraft.

Missile Break-Up–Modes were examined to determine if upon ground impact of a non-propulsive missile, fragment would ricochet and intercept the flight path of the aircraft Determination of the missile break-up was found by first determining the velocity of a missile at ground impact as function of the speed of the carrier aircraft. Each missile station shown in Figure 10 was examined to estimate the loads imposed on the missile structure due to impact velocity. These loads were compared to the maximum capability of the missile stations. Analysis was made on impacts on various types of surfaces, water, wet clay, and hard surface to determine structure impact. Figure 11 shows the forward section of the missile plotted against the structure breakup limits. From these data it was concluded that the forward sections of the missile would separate from the motor case, and that while there would be a scattering of missile fragments, none of the fragments would reach a height of more than about 50 feet.

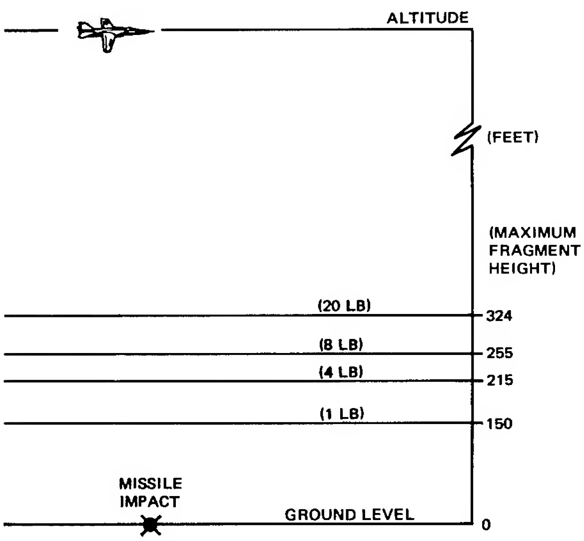


FIGURE 9. MISSILE GROUND IMPACT SAFETY STUDIES

VII. Nuclear Safety Studies

Nuclear Safety analysis was conducted as an integral part of the overall weapon system analysis. The details of these studies cannot be discussed in this paper because of security limitations. However, the studies paralleled and utilized the same analytical techniques that were discussed in section two of this paper. There were seven special nuclear safety studies that were conducted jointly with the payload contractor. These studies covered a two year time period and include a special Crash Safety study to investigate aircraft crash modes with resultant damage to the motor/payload. At the completion of these studies a series of 5 (five) additional nuclear safety analyses were conducted in support of Air Force Weapon Lab (AFWL). These studies were in response to a new Air Force data item which requires the contractor to conduct and document nuclear safety analyses. In addition to the formal studies, the safety organization provided the focal point between The Boeing Co. and AFWL to assure that the design requirements of AFSCM 122-1 (Nuclear Systems Safety Design Manual) were met. The safety group provided technical support to the Nuclear Weapon System Safety Group (NWSSG) in the form of technical briefings on the SRAM system, and technical support on the conduct of the NWSSG's nuclear analysis of the weapon system.

VIII. Analytical Techniques Applied to Production and Operational Program Phase

The production safety program was designed to meet the requirements of MIL-STD-882. The major difference between the DDT&E program and the production/operational program was in the type of analysis conducted by the safety organization and the type of safety activities carried out by the safety group. In general the safety group shifted from a hardware oriented analysis to a people/procedure analysis. The major objective during this phase is to assure that the design features incorporated during the DDT&E were not compromised or that unsafe test and handling practices were not used.

Operating Hazard Analysis–was performed at the missile assembly facility, and at all SRAM bases. The primary objective of this analysis was to examine each step in the assembly, handling, and testing and determine safety standard (requirements) which could be incorporated into the planned engineering control documents for the operations at these facilities. This analysis was completed prior to starting work at the facilities. The second objective of the operating hazard analysis was to provide the safety engineer at each base and the missile assembly a record of the identified hazards for each step in the operation. This could then be utilized by the safety engineer to evaluate possible changes in operating procedures and in evaluating defective equipment.

Safety Surveys and Audits–were conducted on suppliers of safety critical items prior to the start of full production and on a yearly follow-up basis. The objective of these surveys was to assure that the facilities were meeting all the Federal, State, and local safety requirements. Many of these suppliers were single source and it was important to the

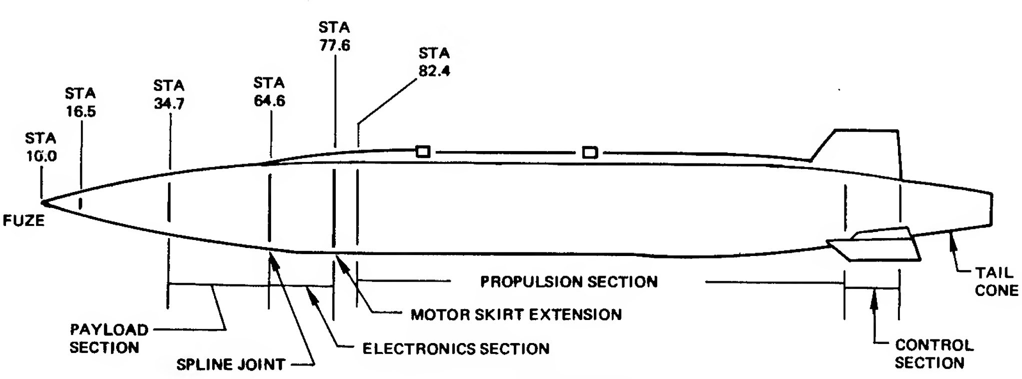


FIGURE 10. MISSILE STATIONS

successful deployment of the SRAM system that these suppliers did not have a serious safety accident which would jeopardize the production schedule. Audits were performed by the safety organization at the missile assembly facility and the SRAM bases to assure that the safety standards identified by the hazard analysis and incorporated into engineering control were being carried out.

Safety Training—was provided as an integral part of all Air Force personnel equipment testing, usage, and maintenance courses. The safety organization received training course material and provided additional course material to the instructor for course preparation.

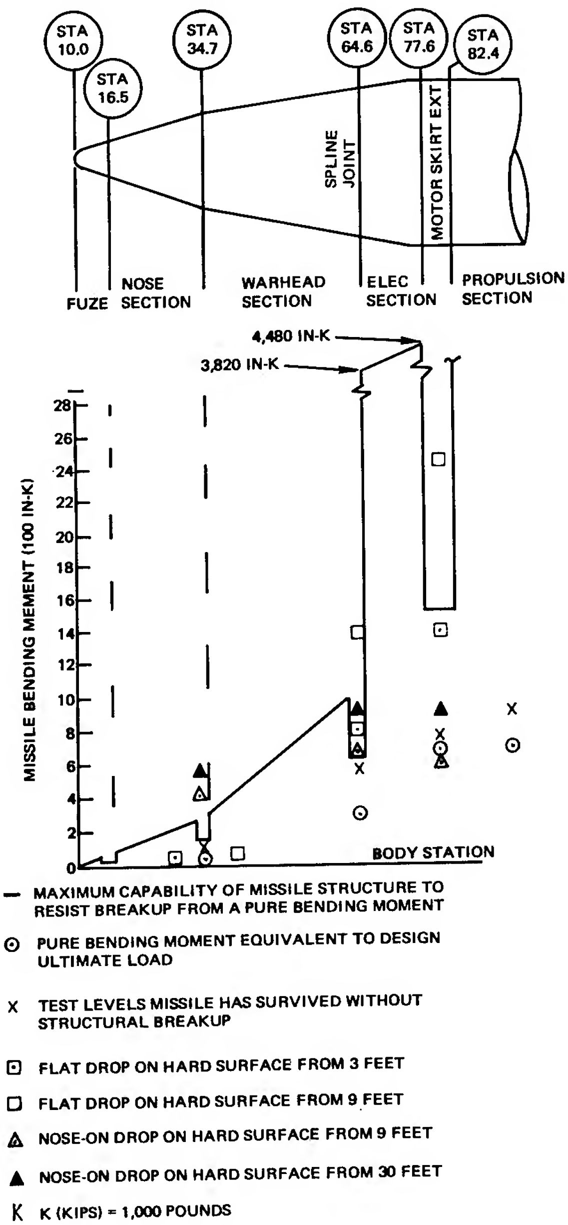


FIGURE 11. MISSILE BREAKUP LIMITS

IX. Conclusions

The application of safety disciplines to the SRAM program represents a unique test for the disciplines that have been developed over the past several years. The SRAM program was one of the first major weapon systems to have system safety considered as an integral part of the program from the very start. System safety is a relatively new discipline, and as such a heavy emphasis has been placed upon the development of analytical techniques to examine safety related problems in a program. These analytical techniques were developed early in the formation of the system safety discipline, but their effectiveness in solving safety problems has been limited due to the ivory tower approach taken by many safety organizations. In other words there has been a lot of talk, many papers, but in some cases it has been difficult to measure the contribution of a safety organization to the program. For system safety to continue to grow as an engineering discipline it is necessary to show that the discipline can contribute actively to the overall success of a program. More emphasis must be placed on positive program accomplishments and less emphasis on arm waving. To accomplish this objective primary emphasis in the safety discipline must be now placed on application of the safety analytical tools that have been under development for the past several years.

The major objective of this paper has been to show that safety analytical tools have been successfully applied to the SRAM program, which has resulted in no accidents or incidents. For the safety organization to assume credit for this record is absurd. As discussed previously the safety organization is a part of the team effort and the credit for this record goes to the total team not just one element. Then how does one measure the effect of safety against the total effect. There are some positive indications which show the safety contribution to the success of a program. These are documented safety analysis, engineering changes which reflect safety as a reason for change, but most important is the recognition of the safety organization by other engineering peer groups. The SRAM safety organization has noted with pleasure that prior to starting of work on design changes, test plans, test procedures the various engineering groups engage the safety engineers in informal discussion to obtain direction.

An important element in the success of the program has been the contribution by the Safety Engineers assigned to the SRAM program who collectively represent a wide range of engineering disciplines. Another important element in the success of the program was the continuous support, direction, and guidance provided by the Air Force Safety manager (Mr. Paul Boyer) of ASD.

The final question is what was the price tag for all this effort, and was it cost effective? The Boeing System Safety program represents approximately 1% of the total effort on the SRAM program. In terms of effectiveness this has to be evaluated in terms of no personnel injury or equipment lost. It is the author's opinion that this small effort is cost effective and that future programs should continue to make safety an active part of the program.

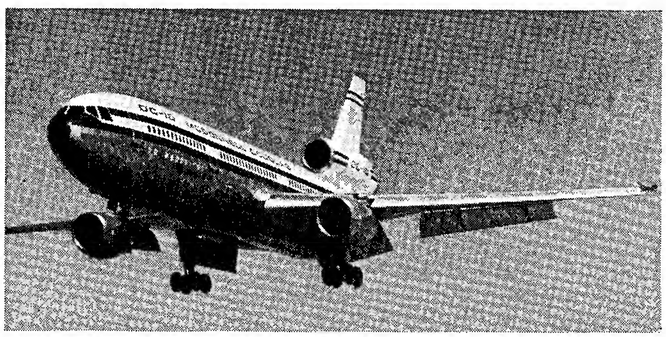
INTEGRATED RELIABILITY AND SAFETY ANALYSIS OF THE DC-10 ALL-WEATHER LANDING SYSTEM

Kenneth L. Peterson, Senior Engineer/Scientist
and
Robert S. Babin, Senior Engineer/Scientist

Reliability and Safety Engineering
Douglas Aircraft Co.
McDonnell Douglas Corp.
3855 Lakewood Boulevard
Long Beach, California

Abstract

The Douglas DC-10 airplane has an all-weather landing (AWL) capability, which permits automatic landings under zero-visibility weather conditions. An exhaustive reliability and safety analysis of the DC-10 AWL System showed conclusively that the safety of such landings, despite any concurrent system failures, is extremely high – higher, in fact, than that of conventional, visual landings under the pilot's control – and that government regulatory agencies' safety criteria are fully met. The system analysis combined several common reliability analysis techniques with extensive modeling of system performance.



Introduction

The Douglas DC-10 is a half-million-pound, commercial aircraft that utilizes three high-thrust jet engines and a wide body to accommodate a large number of passengers while operating from relatively short runways. The size of the DC-10 can be best understood by comparing its profile with that of the world renowned Douglas DC-3 (Figure 1).

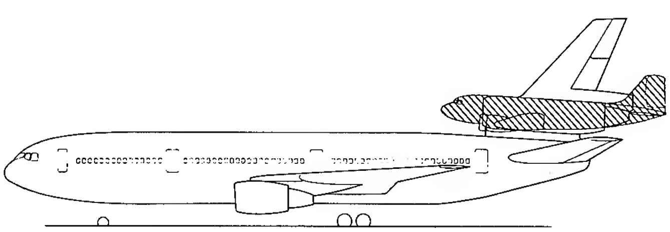


FIGURE 1. SIZE COMPARISON DC-10 VS DC-3

The DC-10 has been in airline service since August 1971 and has established an excellent record for reliability and safety. The broad programs that underlie those achievements are thoroughly described in a recent technical paper ("Development of Douglas Commercial Aircraft Reliability Programs," D. L. Gilles, published in the Proceedings of the 1972 Annual Symposium on Reliability, San Francisco, 25 January 1972).

One of the most significant operating features of the DC-10 is its ability to execute fully automatic landings in all visibility conditions, including the "zero-zero" condition where vertical and horizontal visibility are essentially zero feet. This All-Weather Landing (AWL) capability means that the aircraft can flawlessly execute all required landing maneuvers, regardless of visibility conditions, at any airport equipped with the necessary ground equipment. All-weather landings provide cost savings to airlines and improved service to passengers by greatly reducing the number of bad-weather diversions to alternate airports.

System safety is perhaps the single most significant performance requirement for an AWL system. System safety requirements are in fact so stringent that aircraft using AWL systems built to current government regulatory criteria can be expected to be even safer than visually operated aircraft are today. As an airframe manufacturer, the Douglas Aircraft Company has an obligation to demonstrate that its AWL systems comply with regulatory criteria and qualify for type certification by the U.S. Federal Aviation Administration (FAA) and foreign regulatory agencies. Analyses are required to verify that the systems' performance and reliability are such that for operation in reduced weather minimums they meet the safety requirements of those regulatory agencies. The AWL system must then go through a rigorous training and review cycle with the user airlines before any airline is authorized to use it in revenue operation. The AWL system on the DC-10 will enter this cycle during 1973.

The *DC-10 AWL System* (called "the System" hereafter) is an extension of the conventional automatic-pilot, or "autopilot," system which processes signals from radio receivers, gyroscopes, altimeters, and other sensors to command the aircraft to track ground-transmitted localizer and glideslope beams. Figure 2 shows the spatial relationships of the two beams, the aircraft, and the runway during an automatic landing. (This simplified sketch omits several details; for example, the two beams do not actually originate at the same point on the runway.) The localizer beam, actually the field-intensity pattern of a modulated VHF carrier radiated from a directive ground antenna, is the reference by which the autopilot keeps the aircraft's course coincident with the centerline of the runway. Similarly, the glideslope

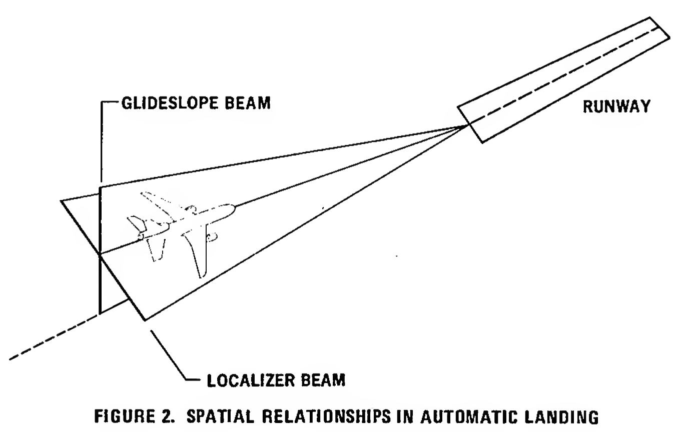


FIGURE 2. SPATIAL RELATIONSHIPS IN AUTOMATIC LANDING

beam, a UHF carrier, serves to hold the aircraft's descent angle to a specified magnitude, roughly three degrees. The System extends the conventional autopilot's capability by achieving completely automatic landings; in contrast, landings with conventional autopilots require the pilot to take over manual control of the aircraft at least several seconds before touchdown.

Designing the System to meet performance and reliability requirements was a difficult task, not only because of the basic system complexity, but also because of the complex airborne-ground systems interfaces and the numerous pilot-aircraft interfaces. These made ordinary numerical reliability analyses inadequate for fully assessing the System's reliability and safety. Therefore a special integration of performance analyses and failure analyses, called the *Reliability and Safety (R and S) Analysis*, was developed for the System. This Analysis utilized data generated by failure mode and effects analyses, fault tree analyses, numerical failure-rate predictions, worst-case circuit analyses, system fault analyses, and special digital-computer simulations, all of which were combined to provide a complete assessment. This integrated analysis provided answers to all questions regarding the effects and annunciation of any fault, single or multiple, its probability of occurrence, and the necessary corrective action by the flight crew following each fault.

AWL System Safety Criteria

The System is designed to comply with the landing safety criteria prescribed by the FAA and the British Civil Aviation Authority (CAA), and it was these criteria that largely established the reliability and safety features of the System design. A detailed description of these criteria is therefore presented to assist in understanding the System and its R and S Analysis.

FAA and CAA landing safety criteria are specified in terms of three landing approach categories, two of which are of concern here (see Figure 3). In a Category II approach, the pilot expects to be able to see the runway when he reaches the landing decision height and can then continue with an automatic landing, or, in the event of a failed landing system can make a manual landing. If, however, he cannot see the runway at decision height, he must immediately execute a go-around. At decision height, since the aircraft is only 100 feet off the ground and only about 13 seconds from touchdown, it is a logical requirement that any landing system failure must not cause the aircraft to make abrupt maneuvers or interfere with the pilot's taking and maintaining control of it. The formal FAA/CAA statements of these requirements, called the fail-passive criteria, are paraphrased in Figure 3. Virtually all American and British jet aircraft in commercial service are currently designed and certified for Category II landing approaches.

Category III requirements characterize all-weather landings, in which the pilot does not need to or expect to see the runway before touchdown. When alert height (in effect, synonymous with "decision height" in Category II) is reached, the pilot must execute a go-around if the landing system is no longer "fail-operational," as defined in Figure 3, which paraphrases the formal FAA/CAA statements of those

LANDING APPROACH CATEGORY	LANDING DECISION (OR ALERT) HEIGHT (FT)	SAFETY CRITERIA
II	100	THE SYSTEM MUST FAIL PASSIVE: ANY FAILURE IMMEDIATELY TELLS THE PILOT HIS SYSTEM HAS FAILED, AND DOES NOT CAUSE ABRUPT AIRCRAFT MANEUVERS OR INTERFERE WITH HIS NORMAL CONTROL OF IT.
III	100	THE SYSTEM MUST FAIL OPERATIONAL: ANY SINGLE FAILURE HAS NO EFFECT ON AIRCRAFT PERFORMANCE, SINCE REDUNDANT AUTOLAND CAPABILITY IS PROVIDED; MULTIPLE-FAULT-CAUSED LOSS OF AUTO-LAND CAPABILITY IS "EXTREMELY IMPROBABLE."

FIGURE 3. FAA/CAA LANDING APPROACH REQUIREMENTS AND SAFETY CRITERIA

criteria. Conversely, if the system is fail-operational at alert height, the automatic landing will proceed. During the remaining seconds before touchdown, the landing will be unaffected by any system failure. Category III landing criteria are an innovation of the late 1960s, and in the United States only the wide-body jet aircraft (DC-10, 747 and L1011) have been designed for such landings.

The failure criteria summarized in Figure 3 constitute the basic reason for the R and S Analysis, most of which was performed to verify the System's compliance with those criteria. A description of the System and the Analysis follows, after some clarifying definitions.

As used herein, a "fault" is a failure of functional output – electrical, mechanical, or hydraulic – of an equipment unit (black box) of the System, which means that the functional output is outside its acceptable limits. The term "fault" also includes failures of inputs to the System from other, interfacing systems on the airplane, for example, voltage from the generators. A "hazardous" fault is one that could conceivably cause an unsafe landing, that is, a landing that fails one or more of the approach and landing safety criteria (see Figure 3). "Multiple fault," as used herein, refers to two or more independent, causally unrelated faults. "Failure," as used herein, carries its conventional meaning: the inability of an item to perform within previously specified limits.

System Description and Modeling

The basic System configuration is shown in Figure 4. It consists of two independent, identical automatic-landing, i.e., "autoland," subsystems, No. 1 shown above and No. 2 below the horizontal dashed line. Thus each autoland is capable of carrying out the required sense, compute, and actuate functions to complete a successful landing, totally unaffected by the existence of any fault in the other autoland. (Although not shown in Figure 4, like flight-control surfaces are mechanically cross-coupled between autolands so that either autoland, or both in unison, can satisfactorily control the surfaces.) Described in reliability engineering terms, the System consists of two parallel-redundant autolands without selective switching. Both autolands are normally operating; the occurrence of a hazardous fault in either one is immediately detected by an on-line comparator, which "disconnects" (in effect, de-energizes) that half of the System, and annunciates the faulted status, all without affecting the other autoland.

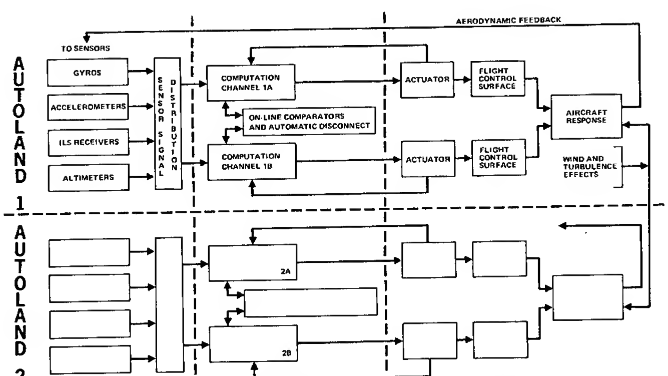


FIGURE 4. BASIC AWL SYSTEM

As shown, the computation function in each autoland is implemented by a pair of identical computation channels: channels 1A and 1B in Autoland 1, and channels 2A and 2B in Autoland 2. The purpose of dual channels is to facilitate fault monitoring, not to increase reliability, since both channels disconnect when either one fails. Each on-line comparator circuit is a differential amplifier whose threshold is set to detect an excessive difference between the two voltages monitored at identical circuit points in the two computational channels. The arrangement permits detection of small, fault-

induced differences between a pair of voltages, both of which normally vary over such a wide range that single-channel fault monitoring would be impracticable. Figure 5 summarizes the vital role of the on-line comparators, without which the rigid safety requirements for the System could not have been met.

44 ON-LINE COMPARATORS (PLUS THEIR ASSOCIATED LOGIC CIRCUITS):

- DETECT EXCESSIVE DIFFERENCE VOLTAGES BETWEEN MONITORED POINTS IN ADJACENT CHANNELS,

- THEN -

- DISCONNECT ONE AUTOLAND (BOTH CHANNELS)

- AND -

- ANNUNCIATE THE FAULTED STATUS OF THE SYSTEM

FIGURE 5. ROLE OF THE ON-LINE FAULT COMPARATORS

Figure 6 is another block diagram of one of the two identical autolands, redrawn to contrast its analog and digital portions. This "AWL half-System" configuration was employed in most of the R and S Analysis, and is therefore called by that name here rather than "autoland." However, the distinction between the System and the half-System is not essential to an understanding of the R and S Analysis, and therefore the term "half-System" will not be used hereafter.

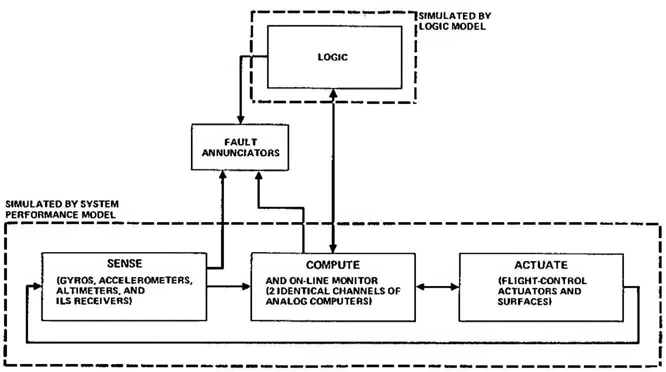


FIGURE 6. RUDIMENTARY AWL HALF-SYSTEM

Within the upper dashed rectangle (Figure 6) are the System's logic circuits, which sense outputs from the on-line monitors and the computers to control the changes in System operational modes and the annunciation of the System's status and its faults. For the R and S Analysis, the operation of this portion of the System was simulated using an IBM 360 digital computer program consisting essentially of the boolean equations that specify the logical functions of these monitoring and annunciation circuits. Called the "Logic Model" hereafter, this simulation permitted a system evaluation of changes in inputs to the logic circuits and of faults within them.

The Fault Annunciators block represents the lights, aural warnings, displays, and other media by which the faulted status of the System is annunciated to the flight crew.

Within the lower dashed rectangle are the sensors, computers, and actuators that carry out those three classic flight-control functions. The computers (two each roll, pitch, and yaw computers in the half-System) are all analog, as are all the electrical signals within this rectangle, with the exception of the digital outputs of the on-line monitors. For the R and S Analysis, all of the functions within this rectangle, that is, all of the sensing, computation, and actuation, plus the associated mechanical systems and aerodynamic feedbacks, were simulated by appropriately programming an IBM 370 digital computer. This computerized simulation model, hereafter referred to as the "System Performance Model," is shown in greater detail in Figure 7. It consists of a digital simulation of the flight control functions with input performance variables and safety criteria, plus the capability (explained later) for inducing specific faults into the Model and determining their effects. Because of its pivitol importance in the R and S Analysis, a more detailed description of that Model follows.

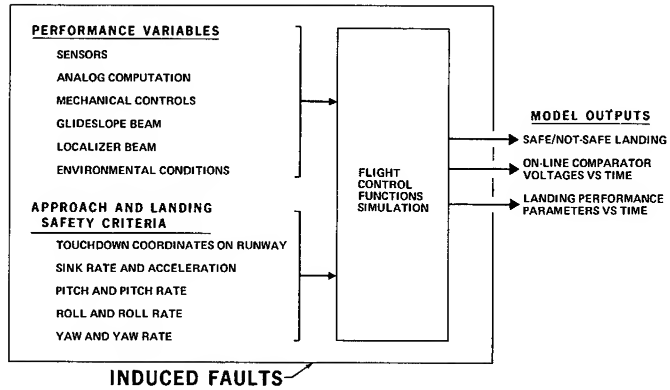


FIGURE 7. SYSTEM PERFORMANCE MODEL

System Performance Model

Effective flight-control functions simulation must accurately represent the actual movement and position of the aircraft with respect to the localizer and glideslope beams and to the ground. To achieve this, all System sensors (Figure 6) were modeled in the way they operate in the aircraft to detect changes in aircraft position. In the compute portion of the simulation, digital representations of analog transfer functions were used to model the actual System circuit operation. To simulate the manner in which the System actually commands the flight control surfaces (Figure 8) to hold the aircraft on the localizer and glideslope beams, the Model was programmed to simulate these commands and the resultant movement of the commanded surfaces and the aircraft by using the six-degree-of-freedom aerodynamic equations of the aircraft. Verification of proper representation of these equations in the Model was established by statistical correlation with data from actual flight test aircraft and from a test fixture called the "Iron Bird," a full-scale DC-10 flight-control mockup with cable runs, hydraulics, actuators, control surfaces and peripheral equipment identical to the production aircraft. The integrated simulation of these sense, compute, and actuate functions was achieved in a closed-loop arrangement in which flight commands, together with changes in sensor outputs, generate commands that feed control-surface actuators, which in turn feed back actuator position information to the computer. The resultant control surface displacement "moved" the simulated aircraft via the aerodynamic equations, thereby satisfying the flight commands.

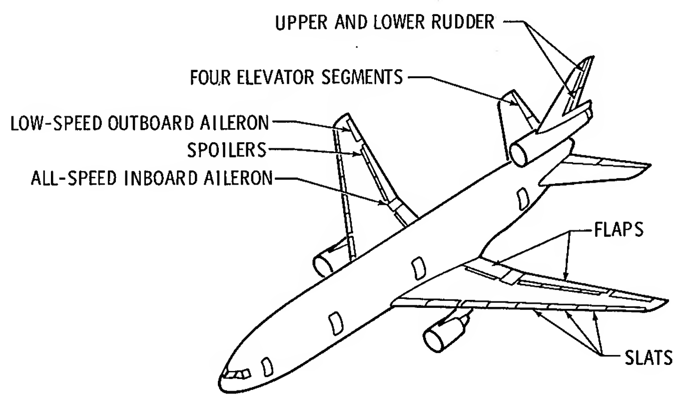


FIGURE 8. DC-10 CONTROL SURFACES SIMULATED IN THE R&S ANALYSIS

Figure 7 lists the System performance variables and safety criteria that are utilized in the flight-control function simulation to "land" the aircraft and assess the safety of the landing. The first three groups of performance variables are the normally expected ranges of values of the key performance parameters of the sensors, analog computer, and mechanical controls (actuators and linkages). Before each simulated landing, a Monte Carlo sampling routine randomly selected as inputs to the Model, a single value for each of those parameters from the distribution of values for that parameter. Each of these distributions had been generated by Monte Carlo sampling of electrical part and mechanical part parameters from their individual distributions. Prior analyses of sizable quantities of data on manufacturing tolerance limits and end-of-life limits for the key parameters of a variety of mechanical and electrical parts, including microcircuits, had shown that the individual part-value distributions were essentially Gaussian.

The statistical distributions of glideslope-beam and localizer-beam variables, namely, their steady-state offsets and variations with time, were provided by FAA. Approximately 1000 times during each simulated landing, the distributions of these two groups of variables were randomly sampled and filtered to limit the rate of change from one sampled value to the next.

The magnitudes of each of the environmental conditions – wind velocity and direction, shear, gusts and turbulence – were randomly sampled every 50 milliseconds during each simulated approach and landing, with filtering to limit the rate of change from one sampled value to the next. The Gaussian distributions describing each of these variables, including the dependence of gusts and shears to wind velocity, were provided by FAA. In the same sense that an aircraft is displaced in real space by these changes in environment, the Model responds by "displacing" the simulated aircraft with respect to glideslope and localizer beams and to the ground.

Acceptable limits for each of the approach and landing safety criteria listed in Figure 7 were determined primarily from analyses and testing of the DC-10 aircraft to determine its structural limitations. Every 50 milliseconds during each simulated approach and landing, aircraft performance was matched against these criteria as part of the flight-control functions simulation. If criteria limits were exceeded at any time, the landing was designated as "not safe."

The System Performance Model provided three categories of outputs for each simulated approach and landing:

1. A binary decision on the safety of the landing, "safe" or "not safe," and a statement of which criteria were violated.

2. A printout of measured difference voltages at each comparator location (or candidate location) versus time during the approach.

3. A quantitative printout of all the landing performance parameters versus time during the approach.

In total, these outputs constitute an overall performance evaluation of the System except for its logic portions.

Reliability and Safety Analysis

Purpose and Prime Goals

Figure 9 summarizes the purpose and prime goals of the Analysis. The 10^{-9} probability number, representing one occurrence in a billion landings, is not explicitly specified by FAA or CAA but approximates the minimum measure of safety those agencies will accept. It can be seen that the R and S Analysis task was essentially one of confirming that all the effects of every conceivable single-fault or multiple-fault occurrence are definitively and quantitatively known and will not prevent a safe landing. Additionally, satisfactory performance of the on-line monitors had to be validated.

BASIC PURPOSE:

TO VERIFY THE AWL SYSTEM'S COMPLIANCE WITH FAA/CAA CRITERIA

PRIME GOALS WERE TO CONFIRM THAT:

- EVERY HAZARDOUS FAULT, SINGLE OR MULTIPLE, WILL BE DETECTED AND ANNUNCIATED AND WILL DISCONNECT THE FAILED SUBSYSTEM
- ANY SINGLE FAULT WILL ALWAYS CAUSE THE AWL SYSTEM TO:
  - A. FAIL PASSIVE IN A CATEGORY II LANDING APPROACH, AND
  - B. FAIL OPERATIONAL IN A CATEGORY III LANDING APPROACH
- THE PROBABILITY OF OCCURRENCE OF ANY COMBINATION OF MULTIPLE FAULTS THAT COULD RESULT IN A VIOLATION OF A OR B IS $<10^{-9}$
- NUISANCE DISCONNECTS CAUSED BY ON-LINE MONITORS RARELY OCCUR.

FIGURE 9. R&S ANALYSIS, PURPOSE AND GOALS

Basic Method and Sequence

The basic R and S Analysis method, summarized in Figure 10, was, in effect, a series of fault effects analyses, but actually consisted of an integrated combination of several kinds of analyses and computer simulation techniques, as will be seen. Figure 11 identifies the four basic phases of the Analysis, preceded by a "fault-free" analysis, which, although not in itself a reliability or safety analysis, was basic to all that followed. The five analyses will be discussed in the sequence shown.

IDENTIFY AND CHARACTERIZE EVERY SYSTEM FAULT

EVALUATE THE EFFECTS OF EVERY SINGLE FAULT:

ON AWL SYSTEM PERFORMANCE

ON AIRPLANE PERFORMANCE

ON THE CREW'S REACTION CAPABILITIES

TO VERIFY ITS MONITORING AND ANNUNCIATION

EVALUATE THE ABOVE EFFECTS FOR EVERY POSSIBLE MULTIPLE-FAULT COMBINATION

EVALUATE SINGLE- AND MULTIPLE-FAULT EFFECTS ON MONITORING AND ANNUNCIATION

FIGURE 10. R&S ANALYSIS, THE BASIC METHOD

FAULT-FREE SYSTEM PERFORMANCE ANALYSIS

FAULT MONITORING ANALYSIS

SINGLE-FAULT EFFECTS ANALYSIS

- ANALOG-FAULT ANALYSIS
- LOGIC-FAULT ANALYSIS

MULTIPLE-FAULT EFFECTS ANALYSIS

INTERFACE-FAULT EFFECTS ANALYSIS

FIGURE 11. R&S ANALYSIS SEQUENCE

Fault-Free System Performance Analysis

As shown in Figure 12, this analysis used the System Performance Model to evaluate the safety level of System performance during several thousand simulated landings. (In this and subsequent flow diagrams, rectangles identify data analysis or computation, and circles and ellipses identify data or information.) During these simulated landings the System was fault free – was operating as designed, with no internal faults – in contrast to its status in the subsequent phases of the R and S Analysis where faults were individually placed in the System before each simulated landing. As shown, corrective action, usually System redesign, eliminated each not-safe condition that was identified. Correlation of the fault-free simulation with actual flight-test data confirmed the accuracy of the System Performance Model and validated its use in the subsequent analyses.

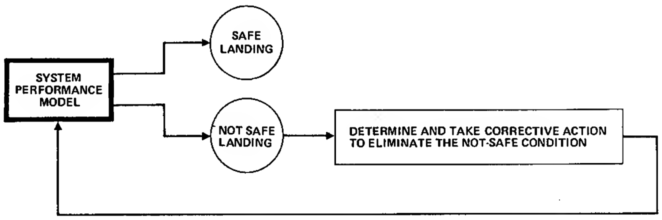


FIGURE 12. PERFORMANCE SIMULATION OF THE FAULT-FREE AWL SYSTEM

Essential Elements of R and S Analysis

A flow diagram of the essential elements of the Analysis (Figure 13) provides an overview of all that is to follow here. As shown, the System Performance Model and Logic Model perform supporting roles to four different analyses: fault-monitoring, single-fault, multiple-fault, and interface-fault. Outputs indicated by the cross-hatched arrow heads are unacceptable or "fail" data that require corrective action. Implementing proper corrective action makes these outputs disappear, and, in effect, the Analysis is satisfactorily completed when all of them have been eliminated. The four different analyses individually described in the paragraphs that follow are interrelated and cannot be viewed as separate, independent entities. However, their partial independence permits a sequential description.

Fault Monitoring Analysis. This analysis, shown in Figure 14, had three goals:

1. To confirm that every hazardous fault is monitored by an on-line comparator (refer to Figure 9).

2. To establish comparator thresholds and delay times to minimize nuisance disconnects while achieving satisfactory fault monitoring.

3. To identify every nonhazardous fault having $Q > 10^{-9}$ (for later use in the multiple-fault analysis).

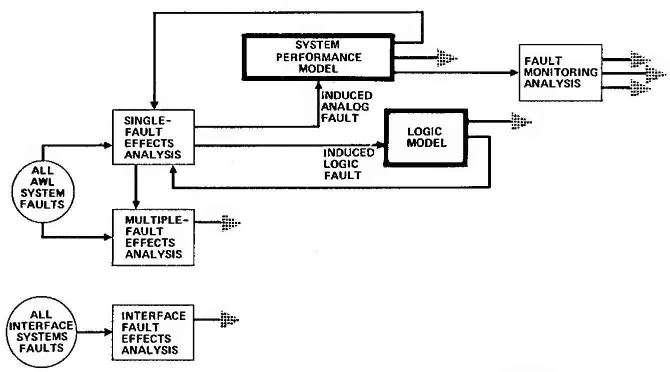


FIGURE 13. ESSENTIAL ELEMENTS OF OVERALL R&S ANALYSIS

As shown, faults (whose nature and derivation will be described shortly) are induced into the System Performance Model one at a time, and the Model's outputs are used to make the monitoring analyses. Achievement of Goal 1 is largely described by Figure 14; when achieved, the "is not" output disappears from all subsequent simulation runs. Nuisance disconnects of on-line comparators (Goal 2) are caused by unequal electrical noise levels in the two monitored channels. Reducing their occurrences to an acceptable number required that the settings of the triggering threshold and response time of each comparator be optimized to diminish its sensitivity to noise while not reducing its required sensitivity to fault occurrences. All potential comparator locations were monitored throughout each simulated landing, and the time-versus-voltage patterns of comparator signals were retained on magnetic tape. Analyses of these signal characteristics from normal, no-fault landings and faulted landings established optimum locations, realistic thresholds, and time delays

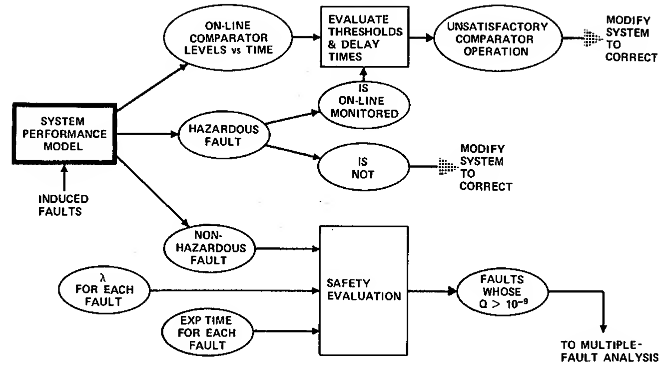


FIGURE 14. FAULT MONITORING ANALYSIS

and locations for on-line comparators to minimize nuisance disconnects.

Achieving Goal 3 required a calculation of Q, the probability of fault occurrence during a landing approach, for every nonhazardous fault, so that faults having $Q > 10^{-9}$ could be used later in the multiple-fault analysis. Q was calculated using the conventional $Q = 1 - e^{-\lambda t}$ formula, where λ is the fault rate and t is the exposure time, i.e., the hours since last tested or monitored and shown not to have failed. Figure 15 puts this calculation into perspective by listing the System's four "levels" of fault detection, in order of increasing exposure time, No. 4 the longest. Some faults are detected only by Level 4 testing, others by Levels 3 and 4, others by all four levels, etc. For every fault, the Q calculation required the determination of maximum exposure time, based on a knowledge of the detection activities at these four levels. λ for each fault was calculated by methods to be described later. Of the four levels, 1 has been discussed in detail, and 3 and 4 are typical of all avionics systems. Level 2 is a pre-land test, executed about 3 minutes before touchdown. It exercises and facilitates monitoring of virtually all of the System, thus limiting exposure time for nearly every fault to 3 minutes.

LEVEL 1: BY ON-LINE COMPARATORS

LEVEL 2: BY PRE-LAND TEST

LEVEL 3: BY UNSCHEDULED REMOVAL ACTION

LEVEL 4: BY ACCEPTANCE TESTING OR OVERHAUL

FIGURE 15. LEVELS OF FAULT DETECTION

Single-Fault Analysis. This analysis had two phases, a Single-Analog-Fault Analysis and a Single-Logic-Fault Analysis, which will now be discussed in that sequence.

Single-Analog-Fault Analysis. Figure 16, the flow diagram for this analysis, shows that here again the System Performance Model was utilized to generate fault monitoring and landing performance

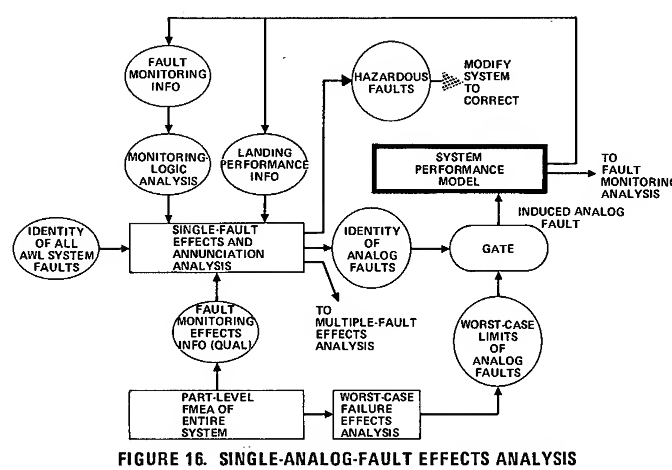


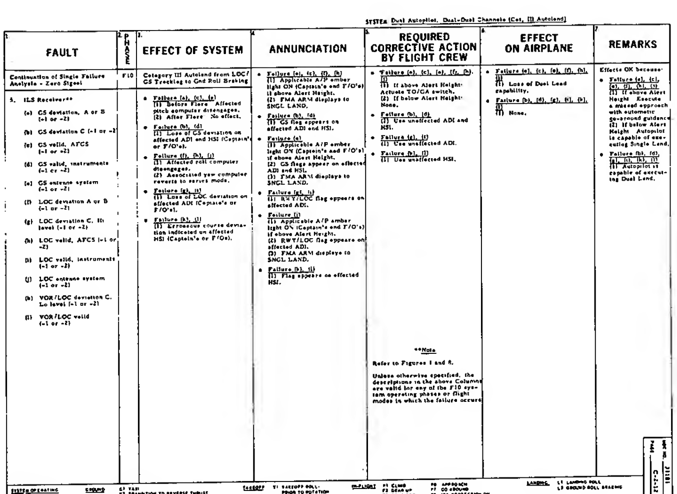
FIGURE 16. SINGLE-ANALOG-FAULT EFFECTS ANALYSIS

information in response to single analog faults induced into it. This analysis had two goals: (1) to eliminate every hazardous single fault, and, after having done so, (2) to develop effects and annunciation data for every single fault, for subsequent use in the Multiple-Faults Effects Analysis.

As one starting point, a basic design analysis established the identity of every fault – analog and digital – in the System, nearly 1600 in total, consisting of 600-odd types of faults, each analog type having three fault modes: maximum possible value, minimum possible value, and a degraded value outside acceptable limits, and each logic type having its two modes.

Next, a conventional Failure Modes and Effects Analysis (FMEA) was made of the entire System, starting at the part level. This overall System FMEA can be viewed as 1600 separate FMEA's, one for each system fault, each deriving the causal relationship between a given fault and the particular group of circuit-part and/or mechanical-part failures that can cause it. As shown in Figure 16, two FMEA outputs are used: (1) fault monitoring effects information is fed to an analysis step that will be explained in the next paragraph, and (2) the identity of each fault and of the particular electrical circuit and/or mechanical assembly that relates to it are fed to a worst-case failure-effects analysis. This analysis produces a pair of quantitative, worst-case limits of every fault – specifically, the worst-case maximum and minimum voltages at every faulted point in the System – as a consequence of the worst single part failure. For the System's electrical circuits, worst-case voltage limits were derived directly; for each mechanical portion of the System, the analog computation circuit simulating that portion was analyzed to derive the voltage limits defining the worst-case fault. These pairs of worst-case analog-fault limits were gated into the System Performance Model, one at a time, as shown, and the Model performed simulated approach and landing runs, each run with the System containing a single fault at its worst-case magnitude. (Note: the induced faults referred to in the Fault Monitoring Analysis, above, are the exact same faults just described.) As shown, the Model's outputs are fed to a Single-Fault Effects and Annunciation Analysis, which will now be described.

The Single-Fault Effects and Annunciation Analysis form shown in Figure 17, one of the 400 such sheets used for this analysis, was used to record the identity of every fault and its effects and annunciation data. For every fault named in the far left column, an analyst, using System schematics, the System Performance Model's outputs with that fault induced, and his evaluation of fault monitoring effects information from the FMEA, entered on the form a statement of the fault's effects on the System (second from left column). In succeeding columns, he entered the manner of its annunciation, the required corrective action by the crew, and its effects on the airplane's capability to continue a satisfactory approach and landing.



SYSTEM Dual Autopilot, Dual-Dual Channels (Cat. III Autoland)

1. FAULT	2. PHASE	3. EFFECT OF SYSTEM	4. ANNUNCIATION	5. REQUIRED CORRECTIVE ACTION BY FLIGHT CREW	6. EFFECT ON AIRPLANE	7. REMARKS
Continuation of Single Failure Analysis – Zero Signal	F10	Category III Autoland from LOC/GS Tracking to Gnd Roll Braking	• Failure (a), (c), (f), (h) (1) Applicable A/P amber light ON (Captain's and F/O's) if above Alert Height. (2) FMA ARM displays to SNGL LAND. • Failure (b), (d) (1) GS flag appears on affected ADI and HSI. • Failure (e) (1) Applicable A/P amber light ON (Captain's and F/O's) if above Alert Height. (2) GS flags appear on affected ADI and HSI. (3) FMA ARM displays to SNGL LAND. • Failure (g), (i) (1) RWY/LOC flag appears on affected ADI. • Failure (j) (1) Applicable A/P amber light ON (Captain's and F/O's) if above Alert Height. (2) RWY/LOC flag appears on affected ADI. (3) FMA ARM displays to SNGL LAND. • Failure (k), (l) (1) Flag appears on affected HSI.	• Failure (a), (c), (e), (f), (h). (1) If above Alert Height: Actuate TO/GA switch. (2) If below Alert Height: None. • Failure (b), (d) (1) Use unaffected ADI and HSI. • Failure (g), (i) (1) Use unaffected ADI. • Failure (k), (l) (1) Use unaffected HSI. **Note Refer to Figures 1 and 8. Unless otherwise specified, the descriptions in the above Columns are valid for any of the F10 system operating phases or flight modes in which the failure occurs.	• Failure (a), (c), (e), (f), (h), (j) (1) Loss of Dual Land capability. • Failure (b), (d), (g), (i), (k), (l) (1) None.	Effects OK because: • Failure (a), (c), (e), (f), (h), (j) (1) If above Alert Height: Execute a missed approach with automatic go-around guidance. (2) If below Alert Height: Autopilot is capable of executing Single Land. • Failure (b), (d), (g), (i), (k), (l) (1) Autopilot is capable of executing Dual Land.
5. ILS Receivers** (a) GS deviation, A or B (-1 or -2) (b) GS deviation C (-1 or -2) (c) GS valid, AFCS (-1 or -2) (d) GS valid, instruments (-1 or -2) (e) GS antenna system (-1 or -2) (f) LOC deviation A or B (-1 or -2) (g) LOC deviation C, Hi level (-1 or -2) (h) LOC valid, AFCS (-1 or -2) (i) LOC valid, instruments (-1 or -2) (j) LOC antenna system (-1 or -2) (k) VOR/LOC deviation C, Lo level (-1 or -2) (l) VOR/LOC valid (-1 or -2)		• Failure (a), (c), (e) (1) Before Flare: Affected pitch computer disengages. (2) After Flare: No effect. • Failure (b), (d) (1) Loss of GS deviation on affected ADI and HSI (Captain's or F/O's). • Failure (f), (h), (j) (1) Affected roll computer disengages. (2) Associated yaw computer reverts to series mode. • Failure (g), (i) (1) Loss of LOC deviation on affected ADI (Captain's or F/O's). • Failure (k), (l) (1) Erroneous course deviation indicated on affected HSI (Captain's or F/Os).				

SYSTEM OPERATING PHASES (COL. 2): GROUND: G1 TAXI; G2 TRANSITION TO REVERSE THRUST; G3 IN REVERSE THRUST; G4 TRANSITION TO STOW & ROLL OUT; G5 AIRPLANE STATIC-SYSTEM OPERATING. TAKEOFF: T1 TAKEOFF ROLL-PRIOR TO ROTATION; T2 TAKEOFF-AFTER ROTATION. IN-FLIGHT: F1 CLIMB; F2 GEAR UP; F3 GEAR DOWN; F4 CRUISE; F5 DESCENT; F6 APPROACH; F7 GO AROUND; F8 ICE PROTECTION ON; F9 BEFORE LANDING; F10 OTHER (DESCRIBE). LANDING: L1 LANDING ROLL; L2 GROUND ROLL BRAKING.

PAGE C-2-12

FIGURE 17. SINGLE-FAULT EFFECTS AND ANNUNCIATION ANALYSIS FORM

To achieve Goal 1, the analyst's review of the effects and annunciation data identified every hazardous fault, all of which were then changed to nonhazardous by appropriate corrective action, for example, by adding an on-line comparator and disconnect, or by increasing the scope of the Pre-Land Test to exercise that function, or by utilizing an alternate sensor to obviate use of that function below alert height, etc. The cross-hatched-arrow output in Figure 16 then disappeared. To achieve Goal 2, the analytical process described above was used to generate the effects and annunciation information for every single analog fault – all of them now nonhazardous – for later use in the Multiple-Fault Effects Analysis.

Single-Logic-Fault Analysis. This analysis had the same goals as the Single-Analog-Fault Analysis and was implemented essentially the same, except that the Logic Model instead of the System Performance Model did the required simulation to determine fault effects and annunciation status. Figure 18 is the flow diagram. All logic faults (selected from the Single-Fault Effects and Annunciation Analysis form) were entered into the Logic Model, one at a time, as an erroneous binary digit. For each entered logic fault, the Model generated information on the annunciation of the fault and on the status of logic circuits affected. Using this information and his System schematics, the analyst: (1) identified hazardous logic faults, all of which were changed to nonhazardous by methods exemplified above, and (2) entered fault effects data on the Single-Fault Effects and Annunciation form for later use in the Multiple-Fault Analysis.

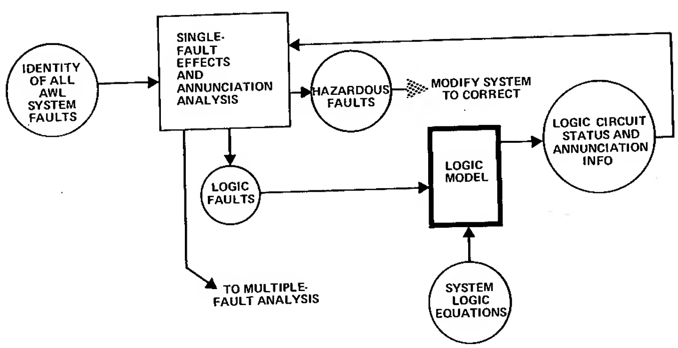


FIGURE 18. SINGLE-LOGIC-FAULT EFFECTS ANALYSIS

Note that the single-fault analyses, just described, established that the System contains no single fault that is hazardous, thus showing that System performance, temporarily ignoring its interfaces, complies with Category II safety criteria (Figure 3). The Multiple-Fault Analysis, to be described next, established that no multiple-fault in the System having a probability of occurrence greater than 10^{-9} is hazardous, thus showing that System performance, temporarily ignoring its interfaces, complies with Category III safety criteria (Figure 3). To be described here last, an Interface Fault Effects Analysis established that no fault or fault combination in interfacing systems could modify either of the preceding statements.

Multiple-Fault Analysis. As shown in Figure 19, a special Fault Matrix served as a vehicle to integrate and document fault and fault-effects data from two of the earlier analyses plus a calculated fault occurrence rate for every fault. Fault rate calculations were made conventionally by combining individual failure rates of parts in each of the circuits and mechanical units analyzed.

Figure 20 shows the format of the Fault Matrix, one of nearly 300 such sheets used in the analysis. The legend at the bottom names the 12 discrete status indicators used to document: (1) the effects of each fault on the pitch, roll, and yaw actuators status and on the commands to them, and (2) the manner and status of the annunciation of each fault.

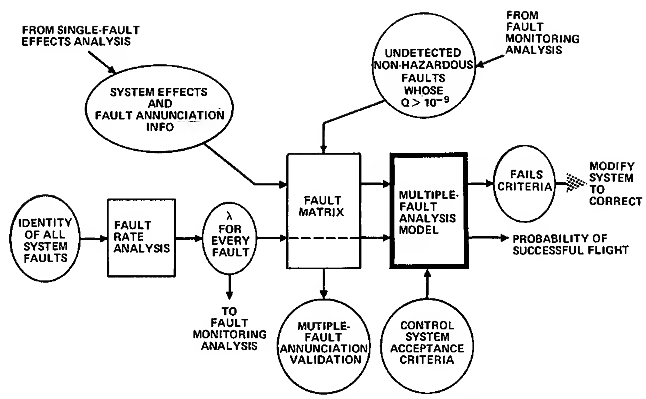


FIGURE 19. MULTIPLE-FAULT EFFECTS ANALYSIS

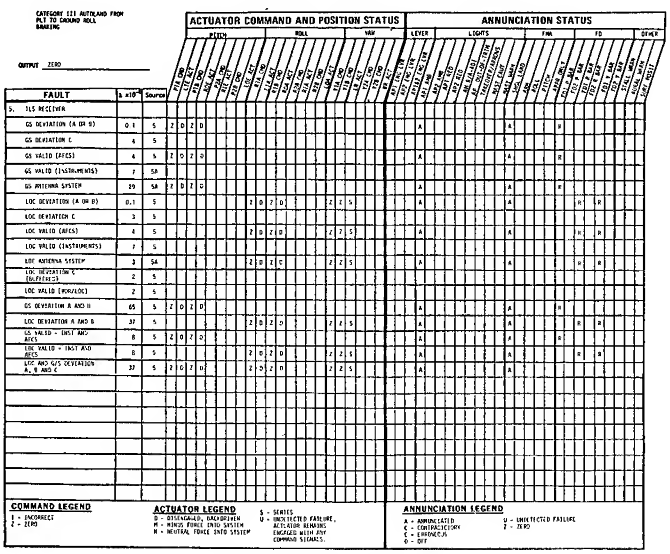


FIGURE 20. FAULT MATRIX

A computer program called the Multiple-Fault Analysis Model (see Figure 19) was written for the IBM 370 computer and was used to:

1. Accept the discretely indicated fault effects (1 above) and the calculated fault-rate quantity for every fault

2. Compare fault effects with stored system performance criteria and thus identify every fault and fault combination that could cause an unsafe landing

3. Use the individual fault rates to calculate the probability of occurrence of every unsafe landing, and

4. Sum the probabilities of all two-fault occurrences that could cause an unsafe landing, and do the same for all three-fault and higher-order fault groupings.

For the probability summations, binomial expansions were used, after certain modifications were made to them to facilitate completion of the summing task within reasonable computer running time. Table 1 is a summary of the results of step 4, above.

The table demonstrates two vital points:

1. As a consequence of any single fault, the probability of an unsafe landing is zero, as is required, and as a consequence of any pair of faults, the probability is less than 10^{-9}, as is required.

2. The consequences of three-fault and higher-order fault groupings can safely be ignored.

TABLE I

SUMMARY OF MULTIPLE-FAULT PROBABILITIES

NUMBER OF INDEPENDENT FAULTS DURING APPROACH, BELOW 100 FT	COMBINED PROBABILITY OF OCCURRENCE OF NUMBER OF FAULTS SHOWN AND OF AN UNSAFE LANDING
1	0
2	6.9×10^{-10}
3 OR MORE	$<5 \times 10^{-12}$

Additionally, the annunciator status indicators tabulated in the Fault Matrix (Figure 20) were reviewed by analysts for every significant pair of faults, i.e., pairs whose $Q_1 Q_2$ product $> 10^{-10}$. That review confirmed that each such multiple-fault occurrence would be immediately detected and annunciated, as is required.

In addition to the calculations already discussed, and beyond the scope of this paper, the Multiple-Fault Analysis Model was also utilized to calculate the probabilities of several events not defined heretofore, i.e., (1) successful go-around: a pilot-initiated maneuver to discontinue the automatic landing and manually fly the aircraft to another landing site; (2) unsuccessful go-around; and (3) being at too low an altitude to safely execute the go-around. These probabilities, calculated using the fault grouping method described above, provided valuable additional insight into the capabilities of the AWL System and of the adequacy of the equipment necessary to execute the go-around.

Interface Fault Effects Analysis. The three airborne systems that interface with the AWL System are the electrical power system, the hydraulic power system, and the ground-sensing system. Conventional Fault Tree Analyses were made to identify and to quantify the probability of every single fault and of every multiple-fault combination in the interfacing systems that, by terminating or degrading a necessary input to the AWL System, could: (1) cause one or both autolands to disconnect; and/or (2) "cause abrupt aircraft maneuvers or interfere with . . . [the pilot's] normal control of it" (Figure 3). Additionally analyzed were the effects of the combined occurrences of a fault in one autoland and a terminated or degraded necessary interface input to the other autoland, the consequence of each such fault pair being the disconnect of both autolands.

Results of this analysis showed that no fault or fault combination in the interfacing systems, or faults in those systems combined with faults in the System, can cause System performance to violate any of its safety criteria (Figure 3).

Conclusions

The final results of the DC-10 R and S Analysis conclusively demonstrated that System performance fully complies with all applicable FAA and CAA safety criteria for Category II and III landings and that nuisance disconnect occurrences are acceptably few. While achieving these results, the Analysis also provided a valuable, running assessment of compliance with criteria as the System design evolved, thus identifying required redesigns to achieve adequate redundancy, to properly locate on-line comparators and set their thresholds, and to decrease part tolerance and drift allowances.

The use of integrated systems analyses, similar in basic concept and implementation to the R and S Analysis, can be expected to increase in the future, as system complexities and safety/reliability requirements grow. The Analysis described above may help illuminate the problems and constraints facing those whose task is the development of unified, efficient analyses of high-performance systems.

MICROCOMPUTERS AND QC. STATISTICS: PROGRAMS AND TABLES*

John H K. Kao**
New York University
Bronx, N. Y. 10453

In this paper, twenty five microcomputer programs are created to solve problems in Quality Control chart work, using Olivetti P-101, Hewlett-Packard 9100B and Wang 700A programmable calculators. As a result, tables for Control Chart constants which are superior to those originally published in ASTM Manual or recently published in ASQC Standard are presented. Numerical examples are given for each area of application.

INTRODUCTION

In recent years as the electronic computers became bigger, faster and more complex, the frustration on the part of the users also grew. Rising from the horizon of scientific computation are three important developments of great significance. They are: time sharing systems, minicomputers and programmable calculators (Some term the last group microcomputers). Perhaps these new developments have actually been induced by or developed in answer to this frustration.

Time sharing provides the user with a "key" to the often inpenetrable, costly but efficient present-day computer system. Software in conversational mode is developed to make the access to a large computer much easier. However, the cost of terminal rental plus fixed revenue such as TCT-terminal connect time, I/O-input output charges over CPU-central process usage time charge and security (or lack of it) of proprietary data and problems are notable shortcomings. Minicomputers then emerged, apparently free of these drawbacks but lacking the interactive quality of the time sharing system. It would be absurd to develop conversational style software for time sharing on a small computer such as the "mini". Being actually a smaller computer, a "mini" acts exactly like its big brother, but at a slower speed (fortunately also at lower cost). As a result, the barrier between the computer and its user remains just about the same. At least one person will have to be in charge of administering a mini computer installation and the user's problem still has to pass through this "administration". The third development, the programmable calculator (microcomputer), probably answered the prayers of the frustrated scientists and engineers. First of all, it is the least costly answer of the three. It is an extended calculator which requires of the user to learn a minimum amount of machine control language in order to program a problem. It could be either shared by many people or be the "private computer" of an individual. Granted, it is limited in capacity, i.e., memory-bound, but the kinds of problems it can handle and the speed with which the answers are obtained rival many current minicomputers. It requires no more space than a desktop. On the other hand, a minicomputer installation, in addition to the main frame, often requires power supply, control panel, input/output peripherals, mounting hardware and altogether needs the space of no less than a good sized (perhaps even air-conditioned) room.

Although this report concerns itself only with the application of programmable calculators, it does not imply that the other two developments viz, time sharing systems and minicomputers, are without merit. In all fairness, we must say that these three developments are complementary to, rather than competing against each other. Together they fill the increasingly widening computational gap left between their big brother - the modern full-scale computer system and the lowly desk calculator - slide rule combinations.

In the pages to follow, we set out to illustrate how the new programmable calculators can actually help solving some very important statistical quality control problems which are "too small" for the modern computer yet "too laborious" for the regular calculator. We adopted the three, to our judgement, most advanced programmable calculators out of a dozen or so available on today's market. These models are: Olivetti P-101, Hewlett-Packard 9100B and Wang 700A*. We created a total of 25 programs (Cf. Appendix) using the languages of these three machines. The first letter of our program identifies for which machine (O-Olivetti, etc.) the program is to be used. None of the 25 programs presented here are available from the published material or manufacturer's software libraries and yet the problems which these programs deal with are of prime importance in the area of applied statistics and quality control. The first 8 programs are for data reduction, i.e. calculating sample moments and other statistics for ungrouped as well as grouped data in such a way as to be most useful in quality control work The next 7 programs generate some 28 constants commonly used in quality control. Two of these constants, viz d_2 and d_3 (for normal distribution), were obtained by double and triple numerical integration on a Digital Equipment PDP-7, and UNIVAC-1108 since none of the present programmable calculators has this capability. The final 10 programs calculate the control limits for 14 control chargs, viz, $\overline{X}$, σ, R, p, np, c and u-charts for both standards known and unknown.

DATA REDUCTION

We shall start with the programs on data reduction. The following definitions using as much as possible ASQC standard A1 [1] for notations are relevant:

(1) $m'_k = \frac{1}{n}\sum^{n} x_i^{\,k} = \frac{1}{n}\sum^{h} x_i^{\,k} f_i$, the sample ungrouped and grouped k^{th} moment about the origin; a special case of this is when $k = 1$, $m'_1 = \overline{X}$, the sample arithmetical mean.

*Work supported by the Office of Naval Research under Contract N00014-67-A-0467-0025 with New York University. Reproduction of this report in whole or in part is permitted for any purpose of the U. S. Government.

**Professor, Department of Industrial Engineering and Operations Research, New York University, Bronx, New York 10453.

*In June, 1971, Olivetti announced their new model P-602 with not only enlarged memory but also improved capability. But the new machine requires a machine language different from that of P-101 and so far no software library is available for P-602. In August 1971, Hewlett-Packard also announced their new model 10, (9800 series) with many improved features, however the machine will not be available until early 1972.

(2) $m_k = \frac{1}{n}\sum^{n}(x_i-\bar{x})^k = \frac{1}{n}\sum^{h}(x_i-\bar{x})^k f_i$, the sample ungrouped and grouped k^{th} moment about $\bar{x}$; a special case of this is when $k = 2$, $m_2 = \sigma^2$, the sample variance. The notation $s^2 = (\frac{n}{n-1})\sigma^2$ will be called the unbiased estimate of variance. σ is the sample standard deviation and s is the root of the unbiased estimate of variance. One can always express m_k in terms of m'_k. As a matter of fact, it is easy to prove that

$$m_k = \sum_{j=0}^{k} \binom{k}{j} m'_{k-j}(-m'_1)^j \text{ so that}$$

$$m_2 = m'_2-(m'_1)^2,\ m_3 = m'_3-3m'_2m'_1 + 2(m'_1)^3 \text{ and}$$

$$m_4 = m'_4 - 4m'_3m'_1 + 6m'_2(m'_1)^2 - 3(m'_1)^4$$

(3) $a_3 = m_3/\sigma^3$, the sample skewness.

(4) $a_4 = m_4/\sigma^4$, the sample peakedness.

(5) $R = x[n] - x[1]$, the sample range where $x[i]$ is the i^{th} order statistic, i.e., $x[1]$ is the smallest and $x[n]$ the largest observation in a subgroup of size n .

(6) $(\text{m.d.})_c = \frac{1}{n}\sum^{n}|x_i-c| = \frac{1}{n}\sum^{h}|x_i-c|f_i$, the sample mean (or average) deviation about c; the two commonly used c-values are $\bar{x}$ and sample median, $\tilde{x}$ which equals $x\left[\frac{n+1}{2}\right]$ if n is odd and equals $\frac{1}{2}(x[\frac{n}{2}] + x[\frac{n}{2}+1])$ if n is even.

The above 6 definitions are elementary statistical notions. Computationally, they are not too laborious to deal with if the values of observations x_i do not differ too much among themselves and the number of observations n is not too large. However, in an industrial setting, oftentimes the data do not behave and usually are massive. As a result, various computational devices such as coding and grouping are often employed to relief the cumbersome and tedious calculations for getting, at least approximate, answers to these statistics. With programmable calculators, these computational devices, although still useful are no longer necessary. We shall illustrate this point by actually using the first 8 programs on a data set of moderately large size. We shall choose the New York City monthly average temperature which is"well-behaved". Perhaps some New York City residents will argue on this, but our choice of the descriptive "well-behaved" has no emotional content. It simply means "statistically stable" which fact will also be verified by the use of our later programs on $\bar{x}$, σ and R-charts. The temperature data which were extracted from a bulletin put out by the Department of Commerce, Environmental Data Service [12] are shown in Table 1. Note the coding opportunity that the data presents. "40" could be subtracted with ease while entering data into the calculator. Now we proceed with our programs whose listings are shown in Appendix. To clear the appropriate working registers, the button marked "Reset", "End" or "Prime" should be depressed on Olivetti, Hewlett-Packard, or Wang calculators respectively before the program is read into the core which is done either by keying in the program codings when machine is in "learn" or "record" mode or by loading the program from pre-recorded magnetic card or tape. To execute the program, the machine should always be put in "Run" mode.

1. Code OK-1: For Computing on Olivetti P-101 machine m'_k , $m_k(k = 1,2,3,4,)$, σ, a_3 and a_4 (non-grouped data)

This program is started by depressing "V". Enter each coded x-value then follow it by "S". Having entered all n x-values, depress "Z". This will cause the machine to print out $m'_1(= \bar{u} = \bar{x} - 40)$, m'_2, m'_3 and m'_4 . Without touching the "Reset" button, load the second side of the program and depress "R,S". The machine will now print out almost instantly $m_1(= 0)$, $m_2(= \sigma^2)$, m_3, m_4, σ, a_3 and a_4 . The following are the results from using this program on data ($n = 70$) in Table 1. (See end of text.)

Biometrika Tables [3] list $\pm$ 5% & $\pm$ 1% percentage points of a_3 as $\pm$ 0.459 and $\pm$ 0.673 for n = 70 [3 p. 183, Table 34B]. A graphical extrapolation on their Table 34C [3 p. 184] gives 4.39, 3.80, 2.34 and 2.20 as the respective upper 1%, upper 5%, lower 5% and lower 1% percentage points of a_4 for n = 70. Apparently the data in Table 1 is approximately normally distributed.

2. Code OK-2: For computing on Olivetti P-101 machine $n, \sum x , \sum x^2 , n\sum x^2 - n(\sum x)^2$, s^2 , s, $\bar{x}$, R and σ .

This program is also started by depressing "V". Enter each coded x-value then follow it by "S". Having entered all n x-values, depress "Y". This will cause the n , $\sum x$ and $\sum x^2$. (This step may be skipped if n, $\sum x$, $\sum x^2$ are not needed.) By depressing "Z", the machine will print out: $n\sum x^2 - (\sum x)^2$, s^2, s, $\bar{x}$, R and σ . Further x-values may be read into the machine (after "Y" and/or "Z") to enlarge the subgroup. To start a new subgroup, depress "D,V" to clear all relevant registers serving as accumulators. Table 2 below shows the results of using this program on data in Table 1. Here 14 subgroups of 5 each are computed and the x-values were again coded by letting $u_i = x_i - 40$.

3. Code HK-2: (Similar to OK-2, but for Hewlett-Packard 9100B machine).

This program may be started by depressing "End", "CNT," - continue. Enter each coded x-value then follow it by "CNT" and the machine will display n, $\sum x$ and $\sum x^2$ in the 3 visible registers - (X), (Y) and (Z) respectively. Having entered all n x-values, depress "Set Flag, CNT". This will cause the machine to display s, s^2 and $\sum(x-\bar{x})^2$. Further depressing "CNT" will produce σ, R and $\bar{X}$ in the 3 registers. If the printer Model 9120A is attached, it will print out the last 2 sets of 3 statistics each with a space between the sets.

4. Code WK-2: (Similar to OK-2 but for Wang 700A machine).

This program can be activated by first reading in the entire block of 5 programs. (Wang Calculator has a larger memory core than either Olivetti or Hewlett-Packard machines.) from prerecorded tape cassette into the machine and then depressing "Prime", and "0002" (Special function key). Enter each coded x-value then follow it by "Go" and the machine will display n and $\bar{x}$ in the two visible (X) and (Y) registers. Depressing "Search, 0" will cause the machine to display σ and σ^2. Depress "Go", the machine will now display $x[1]$ and R. Further depressing "Go" will yield s and s^2 in (X) and (Y) registers. The program may be easily modified to have the answers printed out in a formatted manner with labels and comments on an output writer Model 701.

5. Code OK-3: For Computing on Olivetti P-101 machine the Sample Mean Deviation about an arbitrary constant c from both the computing formula and the defining formula (ungrouped data).

This program is activated by depressing "V" for computing formula or depressing "W" for defining formula for the sample mean deviation about c which is defined before as

$$(\text{m.d.})_c = \frac{1}{n}\sum^{n} |x_i - c| \ .$$

From this is is easy to derive the following computing formula:

$$\text{(6a)} \quad (\text{m.d.})_c = \frac{1}{n}\Big[\sum^{n_3} x - \sum^{n_1} x + (n_1 - n_3)c\Big]$$

where, n_1 = number of observations whose value $< c$,

n_2 = " " " " " $= c$,&

n_3 = " " " " " $> c$.

For $c = \tilde{x}$ (coded), the sample median, further refinements may be introduced to this formula. To execute the program first enter the c-value and then follow it by "S". Enter each coded x-value then follow it by "S." Having entered all n x-values, depress "Z" if started with "V", or "Y" if started with "W". Machine will print $(\text{m.d.})_c$. For the data given in Table 1, the program gives $(\text{m.d.})_{\bar{x}} = 2.1751$. The percentage points for the ratio, g of $(\text{m.d.})_{\bar{x}}$ and s is tabulated as Table 34A in [3, p. 183]. For n = 71, the upper 1%, upper 5%, lower 5%, and lower 1% points are, respectively, 0.8515, 0.8376,0.7607, 0.7430. For data in Table 1, g = 2.1751/2.693 = 0.808 which indicates again that the data in Table 1 is approximately normally distributed.

6. Code HK-3: (Similar to OK-3, but for Hewlett-Packard machine)

This program may be started by depressing "End, CNT". Enter c-value and follow it by "CNT". Enter each coded x-value and follow it by "CNT". Having entered all n x-values, depress "Set flag, CNT". This will cause the machine to display c, n and $(\text{m.d.})_c$ in (X), (Y) and (Z) registers respectively. To use the defining formula, depress "End, Go to 58" before entering c-value. Answers will be printed out if Model 9120A printer is attached.

7. Code WK-3: (Similar to OK-3, but for Wang 700A machine)

Having entered previously the entire block of 5 programs in the core, this program is initiated by indexing the special function "0003". Enter c-value first, then "Go". Enter each coded x-value then follow it by "Go". The machine now display n_1 and n_3 in its (X) and (Y) registers. Having keyed in all x-values, key "Search, 1" and machine will display n and $(\text{m.d.})_c$ in its two registers (X) and (Y). For defining formula, index the special function "0103" before entering c-value. If the output writer Model 701 is attached, this program may be easily modified to have the output printed out.

8. Code OK-4: For Computing on Olivetti P-101 machine $\bar{x}$, m_2, a_3 and a_4 (Grouped data)

In pre-computer time, data reduction by grouping and coding were necessary routines especially when higher moments such as m_3 and m_4 are involved. Nowadays, these techniques, although not necessary, are still welcome in the interest of saving computer time. This program, using grouped and coded data, achieves similar results as does program of Code OK-1 but at about half the program length. Below by Tables 3 and 4 we show two stages of grouping (and coding) the data from Table 1: (See end of text)

Having grouped and coded the data, the k^{th} raw moments,

$$m'_k = \frac{1}{n}\sum_{i=1}^{h} u_i^k f_i \quad \text{are relatively easy to compute.}$$

This program may be started by depressing "V". Next enter d = 1 for one-degree grouping, and d = 2 for two-degree grouping. The lowest coded class mark "-7" and "-6" respectively is then entered followed by f_i-values. Having finished entering all h(= 14 and 7 respectively) f_i-values, depressing "Z" will cause the machine to print out answers for the 4 specified statistics. The following are results of using this program on the grouped data from Tables 3 and 4.

Statistics	For Table 3, One-degree Grouping
$\bar{x}$	-0.1714+46.45 = 46.2786
m_2	7.2849
a_3	-0.0097
a_4	2.5920

Statistics	For Table 4, Two-degree Grouping
$\bar{x}$	0.2857 + 45.95 = 46.2357
m_2	7.5755 +
a_3	0.0711
a_4	2.4468

It should be noted that these statistics are approximate values (due to grouping). However they compare quite favorably with the exact statistics obtained previously by program OK-1 which were: $\bar{x}$ = 46.3286, m_2 = 7.2541, a_3 = 0.0320 and a_4 = 2.5375. Note also that the coarse grouping in two-degree intervals, although easier to calculate gives inferior approximations. A program Code 1.10, on p. 11 of[8] (Olivetti offers an extensive program library for their P-101 [7,8 and 9]) also gives the same coded grouped means of -0.1714 for one-degree grouping and 0.2857 for two-degree grouping and a grouped (m.d.)

$$= \frac{1}{n}\sum^{h} |x_i - \bar{x}| f_i = 2.1624 \text{ and } 2.1878 \text{ respectively}$$

as compared to the exact $(\text{m.d.})_{\bar{x}} = \frac{1}{n}\sum^{n} |x_i - \bar{x}| = 2.1751$ obtained previously by Code OK-3 [cf.Eq.(6)] .

CONTROL CHART CONSTANTS

For the seven programs under this group, we do not need the data in Table 1, because we are dealing with populations characteristics rather than sample statistics. In contrast to sample moments and other statistics given by Eqs. (1) through (6), we need the following corresponding definitions for populations. Below we consider only continuous variable X, and f(x) its density function. Similar expressions may be written down for discrete variable. All we need to do are: replacing integral signs with summation signs and calling f(x) a frequency function, instead of density function.

(7) $\mu_k' = \int_{-\infty}^{+\infty} x^k f(x)dx = EX^k$, the population k^{th} moment about the origin; a special case of this is when k = 1, $\mu_1' = \bar{x}' = EX$, the population arithmetical mean.

(8) $\mu_k = \int_{-\infty}^{+\infty} (x-\bar{x}')^k f(x)dx = E(X-\bar{X}')^k$, the population k^{th} moment about $\bar{x}'$; a special case of this is when k = 2, $\mu_2 = \sigma'^2$ = Var x, the population variance where σ' is the population standard deviation. Similar to m_k in Eq. (2),

$$\mu_k = \sum_{j=0}^{k} \binom{k}{j} \mu'_{k-j}(-\mu'_1)^j$$

(9) $\alpha_3 = \mu_3/\sigma'^3$, the population skewness.

(10) $\alpha_4 = \mu_4/\sigma'^4$, the population peakedness.

(11) When f(x) is equal to zero for $-\infty < x < a$ and $b < x < \infty$, the value (b-a) = S is defined as population range of X and the interval (a,b) is known as the support of f(x).

(12) $(\mu.\delta.)_c = \int_{-\infty}^{+\infty} |x-c| f(x)dx = E|x-c|$ the population mean deviation about c; the two commonly used c-values are $\bar{x}'$ and population median, ξ which is defined by the following equation:

$$\int_{-\infty}^{\xi} f(x)\ dx = \int_{\xi}^{\infty} f(x)\ dx = \frac{1}{2}$$

For example, if X has a normal distribution with mean = $\bar{X}'$ and standard deviation σ' for which (a,b)= $(-\infty,\infty)$ and S = ∞, it is possible to show the following [5, p. 108 ff.]

$$\mu_k = E(x-\bar{x}')^k = \prod^{k/2} (k - 2i + 1)\sigma'^k \text{ , for } k \text{ even;}$$

and $\mu_k = 0$, for k odd. (13)

Two special cases of Eq. (13) are $\mu_3 = 0$, $\mu_4 = 3\sigma'^4$, so that $\alpha_3 = 0$ and $\alpha_4 = 3$. Also for the same normal distribution, the following can be found:

$$(\mu.\delta.)_{\bar{x}'} = \int_{-\infty}^{+\infty} \frac{|x-\bar{x}'|}{\sigma'\sqrt{2\pi}} e^{-(x-\bar{x}')^2/2\sigma'^2} dx$$

$$= \sigma' \sqrt{2/\pi} \text{ ,} \qquad (14)$$

Or the ratio $(\mu.\delta.)_{\bar{x}'}/\sigma' = \sqrt{2/\pi} = 0.79788456...$

Eqs. (13) and (14) form the basis for the tests of normality (or departure from normality)[3, p. 61 and 183] which were carried out earlier under Code OK-1 and Code OK-3.

Now we proceed to the second block of our program.

9. Code OK-5: For computing on Olivetti P-101 c_2, c_3 and $1/\sqrt{2n}$ for n = 2(1)∞ .

When X is normal and sample standard deviation, $\sigma = \sqrt{\sum (x_i-\bar{x})^2/n}$, then $c_2 = E\sigma/\sigma' = \sqrt{2}\ \Gamma\left(\frac{n}{2}\right)/\sqrt{n}\ \Gamma\left(\frac{n-1}{2}\right)$ and $c_3 = \sqrt{\text{Var } \sigma}/\sigma'$

$$= \sqrt{\frac{n-1}{n} - c_2^2} \cong 1/\sqrt{2n} \text{ for large } n.$$

This program utilizes the recursive relationship between two values of c_2 given by the above equation for every other n . Upon depressing "V", the program will print out c_2 , c_3 and $1/\sqrt{2n}$ for all even values of n in succession without limit starting from n = 2 until the machine is either switched off or the reset button touched. For odd n , the printout is activated by depressing "W". Table 5 is the result of using this program which tabulates c_2 , c_3 and $1/\sqrt{2n}$ for n = 2(1) 50. Notice the tendancy that c_3 approaches $1/\sqrt{2n}$. Since a recursive scheme is used in this program, large values of n cannot cause overflow condition.

10. Code OK-6: (Similar to OK-5, but for any assigned n)

The program is initiated by depressing "V". Enter the value n for which c_2 , c_3 and/or $1/\sqrt{2n}$ are desired. The computer determines whether n is even or odd and then chooses the correct branch set forth by OK-5 to evaluate and printout the answers. Since the routine is iterative in nature, it will take more time when n is large. However, the printouts are fairly fast for $n \leq 10$.

11. Code OK-7: For Computing on Olivetti P-101, d_2 , u, $\tilde{d}_2$(max), d_2, u for n = 2(1)∞*.

When X is uniform, d_2 , u = $ER/\sigma' = 2\sqrt{3}\ (n-1)/(n+1)$ and d_3, u = $\sqrt{\text{Var } R}/\sigma' = [24(n-1)/(n+2)(n+1)^2]^{\frac{1}{2}}$.

For the exponential case, $f(x) = e^{-x}$, $x > 0$, (with $\sigma' = 1$), the following can be easily shown:

$$ER = d_2 \text{ , } e = \sum_{j=1}^{n-1} 1/j \text{ and} \qquad (15)$$

$$\text{Var } R = (d_3,e)^2 = \sum_{j=1}^{n-1} 1/j^2 \text{ .} \qquad (16)$$

which can be very easily programmed on a microcomputer. This program will, upon depressing "V", start to print out d_2,u d (max)[cf. OK-9] and d_3 , u for all values

* For an explanation of $\tilde{d}_2$(max), see Code OK-9.

of n in succession beginning at $n = 2$. The output may be terminated by either switching off the machine or touching the reset button. Table 6 shows, among other things such as d_2, e, d_3, e, etc. the above 3 statistics for $n = 2(1) 50$.

12. Code OK-8: (Similar to OK-7, but for any assigned n)

By depressing "V" and enter n, the machine will iterate and printout answers at n^{th} iteration. Again for $n \leq 10$, the printout is quite fast.

13. Code OK-9: For Computing on Olivetti P-101 Plackett's d_2 (max) and $\tilde{d}_2$(max) for $n = 2(1)\infty$

In his paper, Plackett [10] stated that "Populations exist for which d_2 is arbitrarily near to zero, while no population will d_2 exceed the following:

$$d_2(\max) = n\sqrt{\frac{2}{(2n-1)!}\left\{(2n-2)! - [(n-1)!]^2\right\}} \quad (17)$$

Gumbel [4, p. 106] showed an easier proof for the same expression. It can be easily shown that for large n by omitting $[(n-1)!]^2$ (Since it is negligible as compared with $(2n-2)!$) Eq. (47) is approximately equal to,

$$\tilde{d}_2(\max) = \sqrt{n+\tfrac{1}{2}} \quad (18)$$

The factorials in Eq. (17) cannot be calculated in a straight forward manner, as they will cause computer overflow even for relatively small n-values. A recursion formula for Eq. (17) was developed for computing d_2(max). The program may be activated by depressing "V" which will cause the machine to print out d_2(max) and its approximation $\tilde{d}_2$(max) for various n-values in succession beginning at $n = 2$. The printout may be terminated by manual intervention of turning off the machine or touching the reset button. The result, for $n = 2(1) 50$ are incorporated in Table 6. Notice for n as small as 10, d_2(max) and $\tilde{d}_2$(max) are comparable in 3 significant figures.

14. Code OK-10: (Similar to OK-9, but for any assigned n)

By depressing "V" and enter n, the machine will start computation and stops to printout answers when n^{th} iteration is reached.

15. Code OK-11: For Computing on Olivetti P-101 19 Constants for Variables Control Charts, Standard Known and Unknown, $n = 2(1) 25$.

There are ten variables control charts receiving the most attention and usage. Their control limits are: (See end of text).

Among the less popular variables control charts such as median-chart, (m.d.)-chart, etc., are four control charts for "individuals" which are also frequently used. These charts for "individuals" or X-chart are nothing more than four special cases for the first four charts listed above respectively when the subgroup size n is equal to unity. The use of s-chart is not a standard practice in industrial plants, however in engineering statistical research especially in laboratories, s are routinely computed because of the unbiasedness of s^2 as an estimate for σ'^2. Since by definition

$$[\text{cf. Eq. (2)}]\ s = \sigma\sqrt{n/(n-1)}\text{ , we therefore have,}$$

$$Es = \sqrt{n/(n-1)}\ E\sigma = c_2\sqrt{n/(n-1)}\ \sigma' = c_4\sigma' \quad (19)$$

$$\text{and Var } s = \frac{n}{n-1}\text{ Var }\sigma = c_3^2\, n/(n-1)\sigma'^2 = (c_5)^2\, \sigma'^2 \quad (20)$$

Eqs. (19) and (20) shows that s as an estimate for σ' is nevertheless biased with a factor c_4 instead of c_2. That is,

$$Es = c_4\sigma' \quad\text{and}\quad \sqrt{\text{Var } s} = c_5\,\sigma' \quad (21)$$

Naturally,

$$c_5 \cong 1/\sqrt{2n},\ \text{for large n.} \quad (22)$$

The inputs for Code OK-11 are: n, c_2, c_3, d_2 and d_3. The values for c_2 and c_3, for $n = 2(1) 50$, were generated by Code OK-5 without any difficulty. The values for d_2 and d_3 of similar accuracy covering equal range of n are very hard to come by. Tippett [11] used an approximate distribution of R(from a normal distribution) and then used a formula for moments of R from this approximate distribution. By Gaussian quadrature, he obtained d_2 in 5-decimal places for $n = 2(1) 1,000$ and d_3 in 3 and 4-decimal places for only $n = 2(1)20$, 200, 500 and 1,000. We used the exact formulas and through a time consuming adaptive numerical integration procedure obtained, in 8 and rounded to 6-decimal places for both d_2 and d_3 as input to Code OK-11.*

This program may be started by depressing "V" before entering n for which the control chart constants are needed. The machine will print out c_4, $E_1 = 3/c_2$ and $E_3 = 3/c_4$ immediately upon entering c_2 for the same n. When c_3 for the same n is entered next, the machine will print out c_5, A, A_1, A_3, B_1, B_2, B_3, B_4, B_5 and B_6. The value of d_2 is now entered which yields $E_2 = 3/d_2$. Finally d_3 is entered to obtain A_2, D_1, D_2, D_3 and D_4. A total of 19 constants may be had in seconds. Table 7 (a through d) is the result of using this program for $n = 2(1)25$. A portion of this table was given in an ASTM publication [2, p. 115] in 3-decimal with a warning note on the accuracy of the last digit. ASQC in its Standard [1] took this ASTM table and added on values for A_3, c_4, B_5 and B_6 but omitted the doubtful d_3 - values upon which values for D_1, D_2, D_3 and D_4 depended. At end of text are the 3-sigma control limits for the ten variables control charts listed above. These control limits help to explain and define the 19 constants which are output of this program. The basic statistical properties of $\theta = (\bar{x}, \sigma, s, R)$ are its population mean $E\theta$ and population standard deviation $\sqrt{\text{Var }\theta}$ and the 3-sigma control limits assume the form: $E\theta \pm 3\sqrt{\text{Var }\theta}$. The following table summarizes these properties:

* We wish to acknowledge our gratitude to Fred Grossman in Programming these formulas for Digital Equipment PDP-7 and UNIVAC-1108.

θ	$E\theta$	$\sqrt{\text{Var } \theta}$	Unbiased estimate of θ
$\overline{X}$	$\overline{X}'$	$\sigma'/\sqrt{n}$	$\overline{\overline{X}}$
σ	$c_2\sigma'$	$c_3\sigma'$	$\overline{\sigma}/c_2$
s	$c_4\sigma'$	$c_5\sigma'$	$\overline{s}/c_4$
R	$d_2\sigma'$	$d_3\sigma'$	$\overline{R}/d_2$

We shall call these constants: c_2, c_3, c_4, c_5, d_2 and d_3 (all of lower-case letters) basic constants and the others (all of upper-case letters), which depend on the basic constants, derived constants.

CONTROL LIMITS

The final ten programs calculate the control limits for the above control charts (a) through (j), as listed at the end of text under OK-11.

16. Code OK-12: For Computing on Olivetti P-101 Center lines and Control limits for x, σ and R-charts (s-chart, optional), Standard known and Unknown, using Basic constants:

In most instances of applying control charts, especially at the outset, the standard (values for $\overline{X}'$ and σ') is unknown. For this reason, $\overline{X}$-chart is seldom used by itself, but rather it is supported by either σ or R-chart (s-chart may be used in place of σ-chart). However, in process capability studies, since only variability is of concern, either σ or R-chart may be used alone without $\overline{X}$-chart. The first part of OK-12, making use of reduced data: $\overline{X}$, σ and/or R given by OK-2, computes the vector $(\overline{X},\sigma,R)$ and several useful subsets, e.g. $\overline{X}$ and $\overline{\sigma}$, R alone, etc. and stores the elements in the proper registers for further processing. The second part of this program takes off from here and computes the control limits for various combinations of control charts with standard unknown. If later on the standard becomes known, the correct values of $\overline{X}'$ and σ may be inserted by destructively overwritten into the appropriate registers and the program will then compute the control limits with standard known. The program is actually capable of producing the computed control limits as its printout for all 10 control charts (a) through (j) listed above in the last section.

The first part of the program may be started by depressing "V" if the entire vector $(\overline{X}, \overline{\sigma}, \overline{R})$ is wanted. "W" if only $\overline{X}$ and $\overline{R}$ are wanted, "Z" if $\overline{X}$ and $\overline{\sigma}$ is wanted, "CY" if just $\overline{X}$ is wanted, before entering the respective data set: $(\overline{x}_j, \sigma_j, R_j)$, $(\overline{x}_j, R_j)$, $(\overline{x}_j, \sigma_j)$, and $(\overline{x}_j)$. In all these cases, the answers are printed out by the machine upon depressing "Y". If, however, only R or σ is desired, "CY" should be depressed before entering (R_j) or (σ_j) and for these latter cases "CW,Y" and "CZ,Y" are necessary to obtain the respective output of $\overline{R}$ or $\overline{\sigma}$. This part of the program not only prints out $\overline{X}$, $\overline{\sigma}$ and $\overline{R}$, which are center lines of $\overline{X}$, R and σ-charts, but also retains their values in "B", "c" and "b" registers for later use. Of course s may be substituted for $\overline{\sigma}$ if s-chart instead of σ-chart is desired. The second part of the program (on another magnetic card) is also initiated by depressing "V". The values for n, c_2, c_3, d_2, d_3 are keyed in following each entry by "S" (c_4, c_5 should be used if s-chart in lieu of σ-chart is wanted). Right after d_3 is entered, the machine will print out the lower and then the upper control limits for Chart (b), $\overline{\overline{X}} \pm 3\,\overline{\sigma}/c_2\sqrt{n}$. Now, depressing "Z" "Y" and "W" will cause the machine to print out the control limits for Charts (c), (f) and (j), i.e., $\overline{\overline{X}} \pm 3\,\overline{R}/d_2\sqrt{n}$, $(1 \pm 3c_3/c_2)\,\overline{\sigma}$ and $(1 \pm 3\,d_3/d_2)\,\overline{R}$ respectively all for standard unknown. If the standard is known, the values for $\overline{x}'$ may be destructively read into B-register and σ' into both b- and c-registers, whereupon the center lines $c_2\sigma'$ and $d_2\sigma'$ may be printed out manually and a further depression of "Z" "Y" and "W" will yield the printout of limits for Charts (a), (i) and (e), i.e., $\overline{X}' \pm 3\sigma'/\sqrt{n}$, $(d_2 \pm 3d_3)\sigma'$ and $(c_2 \pm 3c_3)\,\sigma'$ respectively. (Control limits for s-chart may be obtained by read in c_4, c_5 in place of c_2, c_3. Then instead of the limits for charts (b) (f) and (e) above we shall have respectively the control limits for Charts (d), (h) and (g) which are: $\overline{\overline{X}} \pm 3\,\overline{s}/c_4\sqrt{n}$, $(1 \pm 3c_5/c_4)\,\overline{s} = (1 \pm 3c_3/c_2)\,\overline{s}$ and $(c_4 \pm 3c_5)\,\sigma'$.)

Table 8 below shows the result of using this program on data summarized by OK-2 in Table 2. The center lines as well as the control limits for all ten control charts: (a) through (j) are calculated in a few minutes. The sample mean of 46.3286 and sample standard deviation of 2.693 for all 14 subgroups are taken as the population standard.

17. Code OK-13: For Computing on Olivetti P-101, Center lines and Control limits for $\overline{X}$, σ and R-charts (s-chart, optional), Standard known and Unknown, using Derived constants.

This program is an alternate for OK-12. Instead of using the basic constants (only 4), it uses the derived constants (there are 19). As a result, the program is much shorter and hence quicker to run. (See end of text)

However, it does require reading in all those derived constants and for good accuracy, does need a table such as our Table 7 (a) through (d) which gives sufficient number of significant places. Using the same data the results of OK-12 and OK-13 on any particular constant will not be exactly the same. This is due to roundings in the derived constants as well as truncations in calculator operations, but they should not be different for the first 4 or 5 decimal places in all cases. We might add, in viewing Tables 2 and 8 together, that none of the 14 $\overline{x}$, σ, s and R-values are outside of their respective control charts (altogether ten in number). Here is a strong indication that the temperature data in Table 1 possess the statistical stability property which we mentioned in the beginning of this paper.

We now take this opportunity to make a few corrections, since we have these new programs and new tables, to the ASTM Manual [2]. The original and corrected results for a few examples taken from the ASTM Manual are shown below in Table 9a (See end of text).

No doubt other examples shown in the ASTM Manual [2] also suffer similar drawback in the absence of reliable tables for control-chart constants and efficient microcomputers which actually reduce the tedious calculations to practically just keying-in the data.

18. Code HK-12: (Similar to OK-12, but for Hewlett-Packard 9100B Machine)

This program may be initiated by "End" and "CNT". Next read in the values for n, c_2, c_3, d_2, d_3 following each with "CNT" (c_4, c_5 may be used in place of c_2, c_3 if s-chart instead of σ-chart is wanted). Next read in the values for $\bar{x}$, σ and R following each with "CNT". Having entered all k sets of values, for averages of these, touch "Set Flag, CNT", the machine will display $\bar{\bar{x}}$, $\bar{\sigma}$, and $\bar{R}$ in its 3 visible registers - (Z), (Y) and (X) respectively. Depressing 4 "CNT" repeatedly will cause the subsequent display of

$[\bar{\bar{x}} \pm 3\bar{\sigma}/c_2\sqrt{n},\ k]$, $[(1 \pm 3c_3/c_2)\bar{\sigma}, k]$, $[\bar{\bar{X}} \pm 3\bar{R}/d_2\sqrt{n}, k]$ and $[(1 \pm 3d_3/d_2\sqrt{n}),\ k]$ with the upper control limit in (Z), lower control limit in (Y) and k in (X). If standard is known, key "Set Flag, CNT" then read in $\bar{x}'$ and σ' following each with "CNT", the machine will then display $[(d_2 \pm 3d_3)\ \sigma',\ k]$ and $[(c_2 \pm 3c_3)\ \sigma', k]$ with the same display format as before. This program will print out all answers if the printer 9120A is attached.

19. Code WK-12: (Similar to OK-12, but for Wang 700A machine)

Index the special function key "0012" after loading the entire block of 5 programs into the core. Next read in the values for n, c_2, c_3, d_2, d_3 following each with "Go" (c_4, c_5 may be used in place of c_2, c_3 if s-chart instead of σ-chart is wanted). Next read in the values for $\bar{x}$, σ and R following each with "Go". Having entered all k sets of values, depress "Search, 2", the machine will display $[k,\bar{\sigma}]$ with k in (X) register and $\bar{\sigma}$ in (Y) registers. Key "Go" and the machine will display $[\bar{R}, \bar{\bar{x}}]$ in similar format. Depressing 4 "Go" repeatedly will cause the subsequent display of $[\bar{\bar{X}} \pm 3\bar{\sigma}/c_2\sqrt{n}]$, $[(1 \pm 3c_3/c_2)\bar{\sigma}]$, $[\bar{\bar{X}} \pm 3\bar{R}/d_2\sqrt{n}]$ and $[(1 \pm 3d_3/d_2)\bar{R}]$ with lower limits in (X) and upper limits in (Y). For standard known, key "Search, 3" and then read in $\bar{x}'$ and σ' following each with "Go", the machine will display $[x' + 3\sigma'/\sqrt{n}]$. Two additional keyings of "Go" yield $[(c_2 \pm 3c_3)\ \sigma']$ and $[(d_2 \pm 3d_3)\sigma']$ in the same display format. Answers may be printed out.

20. Code OK-14: For Computing on Olivetti P-101, Center lines $\bar{p}$ (or u, plotting points p_i (or u_i) and control limits (of varying width) for p (or u)chart, Standard known and Unknown.

For a stable process, its shrinkage (population value) - the proportion of defective items relative to total number of items produced - is a constant, although oftentimes its value may be unknown. It is designated by p'. Sampling from this process with subgroup size (or sample size) n will yield a binomial random variable X representing the number of defectives to be found in the sample with EX = np' and Var X = np'q' where q' = 1-p'. For large n, X has approximately a normal distribution with the same parameters EX and Var X. Therefore the sample proportion defective, p = X/n is also approximately normal for large n with EX/n = p' and Var X/n = p'q'/n. The shrinkage reports of a plant listing the numbers of defectives x_i and corresponding subgroup size n_i, usually are for large n_i, although they may vary from subgroup to subgroup.

(a) p-chart: The p-chart is a plot of these sample proportion defectives $p_i = x_i/n_i$ for i = 1,2,...k with the following control limits:

$$p' \pm 3\sqrt{p'q'/n_i}\ , \text{ for standard known and} \tag{23}$$

$$\bar{p} \pm 3\sqrt{\bar{p}\ \bar{q}/n_i}\ , \text{ for Standard unknown} \tag{24}$$

where p' is the known or aimed-at value and

$$\bar{p} = \sum^{k} x_i/\sum^{k} n_i = \sum^{k} n_ip_i/\sum^{k} n_i\ , \text{ a weighted average} \tag{25}$$

and for all n_i = n, for all i,

$$\bar{p} = \sum^{k} np_i/\sum^{k} n = n\sum^{k} p_i/k(n) = \sum^{k} p_i/k, \text{ a straight average} \tag{26}$$

(b) np-chart: In view of the above, there is no reason for computing p_i, if n_i = n, for all i. Instead, the k x-values (x = np) in the shrinkage report are plotted as a np-chart with the following control limits:

$$np' \pm 3\sqrt{n\ p'q'}\ , \text{ for standard known and} \tag{27}$$

$$n\bar{p} \pm 3\sqrt{n\bar{p}\ \bar{q}}\ , \text{ for standard unknown} \tag{28}$$

with $\bar{p}$ given by Eq. (26) above.

(c) u-chart: When the process is continuous in nature, the shrinkage reports are no longer appropriate. In its place, the process is monitored by the so-called unit-defect report in which the number of defects x_i found from various sample blocks (usually of unequal sizes, n_i) are listed and the number of defects per unit $u_i = x_i/n_i$ calculated to reflect the quality status of the process. Again for a stable process, the true (population) value for the number of defects to appear on blocks of equal size is a constant (perhaps unknown). It is designated by c' = nu'. Unlike p' which lies between 0 and 1, both c' and u' can be any non-negative real number depending on the quality status of the process as well as the size of the sampling block chosen. Sampling from this process for any fixed block size will yield a Poisson random variable X representing the number of defects to be found on the block with EX = c' and Var X = c'. For large values of c', X has approximately a normal distribution with the same parameters EX and Var X. Therefore U = X/n is also approximately normal with Eu = c'/n = u' and Var u = c'/n^2 = u'/n. The u-chart is a plot of $u_i = x_i/n_i$ for i = 1,2,...k with the following control limits:

$$u' \pm 3\sqrt{u'/n}, \quad \text{for standard known and} \tag{29}$$

$$\bar{u} \pm 3\sqrt{\bar{u}/n}\ , \quad \text{for standard unknown} \tag{30}$$

$$\text{where } \bar{u} = \sum^{k} x_i/\sum^{k} n_i = \sum^{k} n_iu_i/\sum^{k} n_i\ , \text{ a weighted average} \tag{31}$$

and for all n_i = n, for all i,

$$\bar{u} = \sum^{k} nu_i/\sum^{k} n = n\sum^{k} u_i/k(n) = \sum^{k} u_i/k\ , \text{ a straight average} \tag{32}$$

(d) c-chart: In view of the above, there is no reason for calculating u_i, if $n_i = n$, for all i. Instead, the k x-values ($x = nu = c$, and $\bar{x} = n\bar{u} = \bar{c}$) taken directly from the report are plotted as a c-chart with the following control limits:

$$c' \pm 3\sqrt{c'} \text{ , for standard known and} \tag{33}$$

$$\bar{c} \pm 3\sqrt{\bar{c}} \text{ , for standard unknown, where}$$

$$\bar{c} = n\bar{u} = \bar{x} \text{ .} \tag{34}$$

It can be seen that the points to be plotted onto both np-chart and c-chart are in fact x-values and the charts therefore should be termed binomial X-chart and Poisson X-chart respectively. Since X-chart was already adopted for plotting "Individuals" [cf. OK-11] , we have avoided these cumbersome names and have chosen the equally descriptive np and c-charts as their names.

This program (OK-14) will deal with two of the above four charts, viz. p and u-charts and another program (OK-15) will later take up the other two charts. In addition to center line and control limits (of varying width), this program is also made to compute each plotting point of p as well as u-charts. The first part of the program evaluate $\bar{p}$ or $\bar{u}$ and if the standard p' or u' is known this part of the program may be omitted. The second part requires the reading in of p' of u' (if standard known) destructively to register wherein $\bar{p}$ or $\bar{u}$ is stored, then $p_i = x_i/n_i$ or $u_i = x_i/n_i$ are printed out for each i along with the lower and upper control limits for that point. The usual depression of "V" will start the first part of the program. The values for (n_i, x_i) are entered for i = 1, 2, ...k following each entry with "S". Having entered all k sets of data, a touch "Z" will cause the machine to printout $\bar{p}$ or $\bar{u}$ and store its value in the proper register for further processing. For the second part, "W" is keyed. If the standard is unknown simply depress "S" before re-entering (n_i, x_i) . The machine will immediately printout p_i or u_i . At this junction, if another "S" is depressed, the machine will printout $\bar{p} \pm 3\sqrt{\bar{p}\,\bar{q}/n_i}$. However, if "Y" is depressed instead of "S", the machine will print out $\bar{u} \pm 3\sqrt{\bar{u}/n_i}$. Whether the point "i" is in control or not can be immediately observed in either case, i.e., p-chart or u-chart. At any later time, when the plant manager wishes to revise $\bar{p}$ or $\bar{u}$ or to try out new aimed-at values for p' or u' , simply depress "W" and enter the new value before entering the next set of (n_i, x_i) . The following Tables 9 and 10 are the result of using this program on two examples in ASTM Manual [2] one for a p-chart and the other for a u-chart.

Of the 31 (= k) subgroups in Table 9, there are only 8 different subgroup sizes (n = 200, 330, 510, 550, 510, 640, 800 and 880). However only 2 sets of control limits (for n = 200 and n = 880) were given in the ASTM [2] which is totally inadequate for judging whether or not each subgroup points is out-of-control.

Subgroup points u_i for i = 1,6,10 and 19 (shown in parenthesis in Table 10) are out-of-control, as were so stated in ASTM Manual [2]. However, of the 20 (= k) subgroups, only 3 sizes (n = 20, 25 and 40) are indicated in the data which is a very unlikely event in actual industrial setting. Ordinarily, more than 3 sets of control limits would have to be computed as they only have in the original ASTM calculations.

21. Code HK-14. (Similar to OK-14, but for Hewlett-Packard 9100B machine)

This program may be started by "End" and "CNT". The values for (n_i, x_i) are entered for i = 1,2,...k following each entry with "CNT". Having entered all k sets of data, key "Set Flag, CNT:, the machine will display $\bar{p}$ (or $\bar{u}$) $\sum x$ and $\sum n$ in its 3 visible registers (X), (Y) and (Z) respectively. If the standard is known, p' (or u') may now be indexed on the keyboard which will replace $\bar{p}$ (or $\bar{u}$) already in (X) register, otherwise the value for $\bar{p}$ (or $\bar{u}$) will prevail for further calculations. "CNT", n_i, "CNT" , x_i are then indexed. At this junction, if a p-chart is needed, just depress "CNT" whereupon the machine will display $p_i = x_i/n_i$, $\bar{p} \pm 3\sqrt{\bar{p}\,\bar{q}/n_i}$ (or $p' \pm 3\sqrt{p'q'/n_i}$ in (X), (Y) and (Z) registers with the lower control limit in (Y) and the upper control limit in (Z). However, if a u-chart is wanted, depress "Set Flag, CNT" for the display of $u_i = x_i/n_i$, $\bar{u} \pm 3\sqrt{\bar{u}/n_i}$ (or $u' \pm 3\sqrt{u'/n_i}$) with similar display format as before. If the printer 9120A is attached all these answers will be printed out with a space between sets.

22. Code WK-14: (Similar to OK-14, but for Wang 700A machine)

Index the special function key "0014" after loading the entire block of programs into the core. Next read in (n_i, x_i) for i = 1,2,...k following each entry with "Go" and machine will display, at each step, $\sum n_i$ and $\sum x_i$ in the two visible registers (X) and (Y) respectively. Having entered all k sets of data, touch "Search 4" will cause the machine to display $\bar{p}$ (or $\bar{u}$) and $\sum n_i$ in (X) and (Y) registers respectively. If the standard is known, now is the time to key in p' (or u') which will replace $\bar{p}$ (or $\bar{u}$) already calculated in (X) register. If the standard is unknown, key "Go" to retain $\bar{p}$ (or $\bar{u}$) . Then re-enter (n_i, x_i) and read $p_i = x_i/n_i$ (or u_i) and $\bar{p}$(or $\bar{u}$) in (X) and (Y) registers at each i . At this junction, if a p-chart is involved, key "Go" and the next display will be $\bar{p} \pm 3\sqrt{\bar{p}\,\bar{q}/n_i}$ with the lower control limit in (X) and the upper control limit in (Y). However, if a u-chart is desired, key "Search, 5" and the machine will display $\bar{u} \pm 3\sqrt{\bar{u}/n_i}$ with the same display format. This program may be easily supplemented by a few instructions to print out all answers in any prescribed format with comments and instructions on Model 701 output writer.

23. Code OK-15: For Computing on Olivetti P-101, Center lines $\bar{c}$ (or $n\bar{p}$), and control limits (of fixed width) for c(or np) chart, Standard known and Unknown.

This program deals with the balanced two of the four attributes control charts introduced earlier under the introductory portion of Code OK-14, viz c and np-charts. For these two control charts the plotting points are data themselves and hence contrary to the case of p or u-chart they need not be computed. To start the program, depress "V". Next one must decide whether c or

np-chart is wanted by entering "1" for c-chart and "n" for np-chart. Then x_i are read in for $i = 1,2,\ldots k$. When all k x-values are in, manual printout of $\sum x_i$, and k is available from (b) and (c) registers for c-chart (or $\sum xi$, k and n from (b), (c) and (B) registers for np-chart), and depressing "Z" will cause the machine to print out $\sum xi$ and two $\bar{c}$'s for c-chart (or $\sum x_i$, $n\bar{p}$ and $\bar{p}$ for np-chart). To obtain the control limits, depress "Y" for c-chart (or "W" for np-chart). Now is the time to read in c' (or n p') for standard known. If the standard is unknown, depressing "S" will retain the calculated $\bar{c}$ (or $n\bar{p}$) for further processing. As a matter of fact, the machine will proceed immediately upon the command "S" to compute and print out $\bar{c} - 3\sqrt{\bar{c}}$ and $\bar{c} + 3\sqrt{c}$ or $n\bar{p} - 3\sqrt{n\bar{p}\bar{q}}$ and $n\bar{p} + 3\sqrt{n\bar{p}\bar{q}}$ for p-chart).

The following are two examples, one each for c and np-charts in ASTM Manual [2] using OK-15. For the c-chart, we obtained for Example 11(2) [2, p. 88], $\sum x = 187$, $k = 30$, $\bar{c} = \sum x/k = 2.3$, the lower control limit = 0 and the upper control limit = 13.723327. For the np-chart, we obtained for Example 7 [2, p. 84], $\sum x = 33$, $k = 15$, $n = 400$, $n\bar{p} = \sum x/k = 2.2000$, $\bar{p} = \sum x/nk = 0.0055$, the lower control limit = 0, the upper control limit = 6.637465.

24. Code HK-15: (Similar to OK-15, but for Hewlett-Packard 9100 B machine)

This program may be started by depressing "End" and "CNT". Key "1" for c-chart (or key "n" for $n\bar{p}$ -chart). Enter x_i for $i = 1,2,\ldots k$ following each entry with "CNT". Having entered all k x-values, key "Set Flag, CNT" and read $\bar{c}$, k, $\bar{c}$ (or $n\bar{p}$, k, $\bar{p}$) from (X), (Y), (Z) registers. If the standard is known, key in c'(or np') to replace c(or $n\bar{p}$) now in (X) register, otherwise skip this step. Key "Set Flag, CNT" for c-chart (or simply key "CNT" for np-chart, to obtain k, $c - 3\sqrt{c}$, $\bar{c} + 3\sqrt{c}$ (or k, $n\bar{p}-3\sqrt{n p q}$, $n p + 3\sqrt{n p q}$) in the 3 visible registers (X), (Y) and (Z) respectively. This program will print out all answers when the printer model 9120A is attached.

25. Code WK-15: (Similar to OK-15, but for Wang 700A machine)

Start the program by touching the special function key "0015" after loading the entire block of programs into the core. Key "1" for c-chart (or key "n" for np-chart). Enter x_i for i, 1,2,...k following each entry with "Go", and read k and $\sum x$ in (X) and (Y) registers respectively. Depress "Search, 6" and read $\bar{c}$ (or $n\bar{p}$) and k in (X) and (Y). If the standard is known, c' (or n p') may be read in to replace c (or $n\bar{p}$) already in (X) register. Key "Go" and read 1 (or n) in (X) and $\bar{c}$(or $\bar{p}$) [c' (or p') for standard known] in (Y). For display of lower control limit in (X) and upper control limit in (Y), depress "Search, 7" for c-chart (or depress "Go" for np-chart). Printing option is available with the output writer Model 701 with minor changes to this program.

CONCLUDING REMARKS

We have shown above some of the most important quality control problems solved easily by properly applying the microcomputers - a new breed of programmable desk calculators. Some of these problems such as c_2 and Plackett's d_2(max) [cf. Eq. (17)] which require numerous iterations are obviously impractical for solving on regular desk calculators. Nevertheless, with simple algorithms they are "too small" for efficient use of full scale computers. As a result, they never got solved. Other problems such as reduction of moderately sized data sets can be handled either by desk calculators or, at the other extreme, by full scale computers, but both incur considerable expense and waste. On one hand, they require trained desk calculator operators. (Roomful of desk calculators manned by operators should be a thing of the past.) On the other, for inputing the computer, data-every piece of it need to be punched on cards used once and discarded. Finally, there are problems such as getting the limits of control charts which involve calculations of simple arithmetic but usually of numerous quantity as to render calculations by desk calculators too tedious and calculations by computers too wasteful.

We have demonstrated, with a good programmable calculator of adequate speed and storage capacity, programs such as those presented in this report may be prepared, debugged, recorded and filed. When called upon these programs may be run quickly and problems solved in a matter of minutes. No doubt, in other areas such as statistical teaching and research there must exist similar "small" problems which can also be profitably transferred to the realm of microcomputers.

REFERENCES

1. ASQC Standard A1: Definitions, Symbols, Formulas and Tables for Control Charts, (adopted in August, 1968 superseding ASQC Standard A1-1951.) Reprinted in Quality Progress vol. II, no. 10, (October, 1969), pp. 21-25, also Vol. III, no. 8, pp. 38-41.

2. ASTM Manual on Quality Control of Materials, American Society for Testing and Materials, Philadelphia, Pa., January, 1951.

3. Biometrika Tables for Statisticians, vol. 1 (2nd Edition) Cambridge University, Cambridge (England) 1958.

4. Gumbel, E. J., Statistics of Extremes, Columbia University Press, 1958, pp. 106-108.

5. Harris, Bernard, Theory of Probability, Addison-Wesley, Reading, Mass. 1966, p. 108 ff.

6. Kao, John H. K., "The Beta Distribution in Reliability and Quality Control", Proceedings of Seventh National Symposium on Reliability and Quality Control, 1961, pp. 496-511.

7. Olivetti P-101 Program Library, Mathematical Subroutines, Olivetti Underwood Corporation, March, 1969.

8. Olivetti P-101 Program Library, Statistical Analysis, Olivetti Underwood Corporation, May, 1969, (Code 1.10, pp.9-13).

9. Olivetti P-101 Program Library, Numerical Analysis, Olivetti Underwood Corporation, July, 1969.

10. Plackett, R. L., "Limits of the Ratio of Mean Range to Standard Deviation", Biometrika vol. 34 (1947), pp. 120-122.

11. Tippett, L. H. C., "On Extreme Individuals and the Range of Samples from a Normal Distribution", Biometrika, vol. 7(1925), pp. 368-370.

12. U. S. Weather Service, Local Climatological Data-New York, N. Y. (Central Park), Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Data Service, 1957-1970.

APPENDIX

PROGRAM LISTINGS

Code OK-1: Olivetti P-101 Codings for computing m'_k and m_k for $k = 1(1)4$, s, a_3 & a_4 .

```
Card 1 of 2: AV, a↑, d+, ↓, S, X, A↓↑, /Y, C↓↑, a↑,
d↓, +, C↓↑, V, aY, R↓, e↓↑, ↓, AY, c↓↑, B↓↑, b↓↑, D↓↑,
+, c↓↑, ↓↑, e↓↑, -, A↓↑, /V, C↓↑, a↑, d↓, +, C↓↑, V,
aV, R↓, e↓↑, X, ↓↑, Y, AZ, D↓, C÷, b↓↑, ÷, B↓↑, ÷,
c↓↑, ÷, D↓↑, b□, B□, c□, D□, b↓, AX, d↓↑, RS.
```

```
Card 2 of 2: A*, A+, R□, B↓, d-, A□, e↓↑, B↓, A+, +,
d-, -, bX, ↓↑, c↓, -, A□, C↓↑, B↓, A+, d-, bX, A+, +,
c↓↑, A+, A+, c-, bX, D↓↑, D-, A□, b↓↑, e↓, A√, A□,
eX, C↓↑, C÷, A□, e↓, AX, b↓↑, b÷, A□.
```

Code OK-2: Olivetti P-101 Codings for Computing n, ΣX, ΣX^2, $n\Sigma X^2 - (\Sigma X)^2$, s^2, s, $\bar{X}$, R & σ.

```
AV, S, d↑, c↑, b↑, BV, C↓, b+, C↓↑, ↓, X, B+, B↓↑, D↓,
a↑, d↓, +, D↓↑, S, b↑, d↓, -, /V, ↓↑, d↓↑, aV, b↓, c-,
/W, c+, c↓↑, aW, CV, AY, D□, C□, B□, V, AZ, D↓, AX, -,
b↓↑, C↓, CX, e↓↑, B↓, DX, e-, A□, e↓↑, e↓, b÷, A□, A√,
A□, C↓, D÷, A□, d↓, c-, A□, e↓, A√, D÷, A□, AW, S, C↓,
-, C↓↑, ↓, X, B↓↑, B-, B↓↑, D↓, a↑, d↓, -, D↓↑, /□, CW,
S, S, S, S, S, S, S, S, BW, V, E√, DW, S, S, S, S, S,
S, S, EW, B*, C*, D*, RW, S, S, S, S, S, S, S, S, FW,
c*, d*, V.
```

Code OK-3: Olivetti P-101 Codings for computing Mean Deviation about C, both Defining and Computing Formulas (C may be sample mean or median)

```
AV, C*, c*, D*, d*, e*, S, B↑, /□, BV, S, b↑, D↓, a↑,
d↓, +, D↓↑, b↓, B-, /V, A↓↑, /W, CV, aW, c↓, a↑, d↓,
+, c↓↑, e↓↑, b÷, e↓↑, CV, aV, C↓, a↑, d↓, +, C↓↑, d↓,
b+, d↓↑, CV, AZ, c↓, C-, BX, d+, e-, D÷, A□, /□, V,
AW, D*, C*, S, B↑, /□, EV, S, ↓, B-, A↓↑, D+, D↓↑, C↓,
a↑, d↓, +, C↓↑, DV, AY, D↓, C÷, A□, /□, W.
```

Code OK-4: Olivetti P-101 Codings for computing $\bar{X}$, m_2, a_3 & a_4 using Grouped Data.

```
AV, S, d↑, S, E↑, /□, BV, S, e↑, E↓, eX, c↓↑, c+, c↓↑,
EX, C↓↑, C+, C↓↑, EX, b↓↑, b+, b↓↑, EX, B+, B↓↑, D↓,
e+, D↓↑, E↓, d+, E↓↑, CV, AZ, c↓, D÷, C↓↑, ÷, b↓↑, ÷,
B↓↑, ÷, c↓↑, C□, C↓, AX, d↓↑, b↓, d-, A□, D↓↑, D√, DX,
E↓↑, b↓, A+, +, d-, -, CX, ↓↑, B↓, -, E÷, A□, b↓, A+,
d-, CX, A+, +, B↓↑, CZ, S, S,S,S,S,S,S,S,S,S,S,S,S,S,
S,S, BZ, A+, A+, B-, CX, c↓↑, Y, S,S,S,S,S,S,S,AY,c-,
D÷, D÷, A□ .
```

Code OK-5: Olivetti P-101 Codings for computing c_2, c_3, &$1/\sqrt{2n}$ for $n = 2(1)\infty$.

```
AV, a↑, d↑, b↑, B↑, d□, BV, d↓, AX, C↓↑, B↓, E-, B÷,
C-, A√, A□, B↓, B+, C↓↑, E↓, C÷, A√, A□, B↓, b+, C↓↑,
B↓, C÷, A√, C↓↑, B↓, a↑, dS, +, R□, /□, B↓, E-, C↓↑,
B↓, CX, c÷, C↓↑, d↓, CX, A□, d↓↑, B↓, b+, B↓↑, CV, AW,
a↑, d↑, b↑, a↑, d↓↑, B↑, D↓, d↓↑, d□, CV, Stored Const-
ants: 0.564189584d↑, 0.723601255D↑, 1.0E↑.
```

Code OK-6: Olivetti P-101 Codings for computing c_2 c_3, & $1/\sqrt{2n}$ for any given n.

```
AV, S, b↑, D↓, d↓↑, b↓, a↑, d↑, ÷, /↓↑, X, /V, Y, aV,
a↑, d↓↑, B↑, d↓, a↑, dX, ÷, A√, d↓↑, CV, AY, a↑, d↑,
B↑, d↓, A√, a↑, d↓, ↓↑, ÷, d↓↑, BV, d↓, AX, C↓↑, B↓,
a↑, d↓, -, B÷, C-, A√, C↓↑, b↓, B-, /W, d□, C□, B↓, A+,
A√, a↑, d↓, ↓↑, ÷, A□, /□, V, aW, B↓, a↑, d↑, +, C↓↑,
B↓, C÷, A√, C↓↑, B↓, a↑, d↓, -, c↓↑, B↓, CX, c÷, C↓↑,
d↓, CX, d↓↑, B↓, a↑, d↑, +, B↓↑, CV, Stored Constants:
3.14159265 4D↑.
```

Code OK-7: Olivetti P-101 Codings for computing $d_{2,u}$ $\tilde{d}_2$(max) & $d_{3,u}$ for $n=2(1)\infty$.

```
AV, a↑, d↓, e↑, F↑, a↑, d↓↑, E↑, ↓, F+, f↓↑, BV, e↓,
dX, E÷, A□, e↓, F+, a↑, R-, dS, +, A√, A□, e↓, f÷,
A√, DX, E÷, A□, e↓, F+, R□, /□, e↓↑, E↓, F+, E↓↑, f↓,
F+, f↓↑, CV, Stored Constants: 3.464101615d↑,
4.898979486D↑ .
```

Code OK-8: Olivetti P-101 Codings for computing $d_{2,u}$ d_2(max) & $\tilde{d}_{3,u}$ for any given n .

```
AV, S, C↑, ↓, e+, A√, A□, C↓, E-, B↓↑, C↓, E+, c↓↑,
B↓, dX, c÷, A□, c↓, E+, b↓↑, B↓, b÷, A√, DX, c÷, A□,
/□, V, Stored Constants: 4.898979486D↑, 3.464101615d↑,
1E↑, 0.5e↑.
```

Code OK-9: Olivetti P-101 Codings for computing Plackett's d_2(max) & $\tilde{d}_2$(max) for $n=2(1)\infty$.

```
AV, a↑, d↑, b↑, a↑, d↓, ↓, a↑, d↓↑, ÷, C↓↑, BV, C↓, A+,
b÷, B↓↑, b↓, A+, a↑, d↓, -, a↑, d↑, ↓↑, ÷, B-, A√, bX,
A□, b↓, A+, a↑, d↓, +, A+, B↓↑, b↓, a↑, d↓, +, CX, B÷,
C↓↑, b↓, a↑, R-, dS, +, A√, A□, b↓, a↑, dS, +, R□, /□,
b↓, a↑, d↓, +, b↓↑, CV.
```

Code OK-10: Olivetti P-101 Codings for computing Plackett's d_2(max) & $\tilde{d}_2$(max), any n.

```
AV, S, c↑, a↑, d↑, b↑, a↑, d↓, ↓, a↑, d↓↑, ÷, C↓↑, BV,
C↓, A+, b÷, B↓↑, b↓, A+, a↑, d↓, -, a↑, d↑, ↓↑, ÷, B-,
A√, bX, D↓↑, b↓, a↑, R-, dS, +, A√, d↓↑, c↓, b-, /V,
D□, d□, /□, V, aV, b↓, A+, a↑, d↓, +, A+, B↓↑, b↓, a↑,
d↓, +, CX, B÷, C↓↑, b↓, a↑, d↓, +, b↓↑, CV.
```

Code OK-11: Olivetti P101 Codings for Computing Nineteen Control Chart Constants: A, A_1, A_2, A_3, E_1, E_2, E_3, B_1, B_2, B_3, B_4, D_1, D_2, D_3, D_4, c_4 and c_5, B_5, B_6 .

```
AV, S, d↑, B↑, BV, S, c↑, B↓, a↑, d↓, -, B↓↑, B÷, A√,
B↓↑, c↓, BX, A□, C↓↑, c↓, a↑, d↓↑, ↓↑, ÷, A□, C↓, a↑,
d↓↑, ↓↑, ÷, A□, S, D↑, B↓, X, A□, d↓, A√, a↑, d↓↑, ↓↑,
÷, A□, b↓↑, b↓, c÷, A□, b↓, C÷, A□, D↓, a↑, d↓↑, X, ↓↑,
c↓, -, A□, C↓↑, C↓, +, +, A□, b↓↑, D↓, a↑, d↓↑, X, c÷,
a↑, d↓, ↓↑, -, A□, +, +, A□, C↓, BX, A□, b↓, BX, A□,
CV.
```

Code OK-12: Olivetti P101 Codings for computing $\bar{\bar{X}}$, σ and $\bar{R}$. Control Limits for $\bar{X}$, and R-charts for Standard Known & Unknown, using basic constants.

Card 1 of 2: Center Lines. AV, BV, S, ↓, B+, B↓↑, S, ↓, c+, c↓↑, S, ↓, b+, b↓↑, C↓, d+, C↓↑, CV, AW, EV, S, ↓, B+, B↓↑, S, ↓, b+, b↓↑, C↓, d+, C↓↑, DV, AZ, FV, S, ↓, B+, B↓↑, S, ↓, c+, c↓↑, C↓, d+, C↓↑, RV, BY, EY, S, ↓, B+, B↓↑, C↓, d+, C↓↑, DY, BW, B↓, b↓↑, B*, DY, BZ, B↓, c↓↑, B*, DY, AY, B↓, C÷, B↓↑, B□, c↓, C÷, c↓↑, c□, b↓, C÷, b↓↑, b□, V, EZ, B*, b*, C*, c*, V, Stored Constant: 1.0 d↑ .

Card 2 of 2: Control Limits. FY, R*, RS, BV, C↓, cX, D↓↑, B↓, D-, A□, +, +, A□, S, AZ, b↓, c↓↑, b↓↑, CV, AY, EV, F↓, A+, +, D↓↑, e↓, D+, d↓↑, d↓, -, -, bX, A□, d↓, X, A□, S, AW, c↓, b↓↑, c↓↑, E↓, F↓↑, E↓↑, f↓, e↓↑, f↓↑, DV, S....(48 altogether)....S, AV, S, ↓, A√, a↑, d↓↑, ↓↑, ÷, C↓↑, S, e↑, S, F↑, S, f↑, S, E↑, c↓, e÷, c↓↑, b↓, f÷, b↓↑, RY.

Code OK-13: Olivetti P-101 Codings for computing Control Limits for $\bar{X}$, σ and R-charts for Standard Known & Unknown, using derived Constants.

AV, S, C↑, BV, C↓, cX, F↓↑, B↓, F-, A□, +, +, A□, S, C↑, b↓, c↓↑, CV, AZ, S, B↑, S, c↑, S, C↑, CV, AW, S, c↑, S, F↑, S, E↑, c↓, FX, A□, c↓, EX, A□, W.

Code OK-14: Olivetti P-101 Codings for computing Control Limits for p & u-charts for Standard Known and Unknown.

AV, b*, c*, BV, /□, S, ↓, b+, b↓↑, S, ↓↑, c+, c↓↑, CV, AZ, c↓, b÷, A□, C↓↑, V, AW, S, C↑, EV, S, B↑, ↓, S, ↓↑, ÷, A□, C↓, S, C÷, -, X, AY, B÷, A√, a↑, d↓↑, X, B↓↑, C↓, B-, A□, +, +, A□, /□, DV.

Code OK-15: Olivetti P101 Codings for Computing Control Limits for np and c-charts for Standard Known and Unknown.

AV, b*, c*, S, B↑, BV, S, ↓, b+, b↓↑, c↓, a↑, d↓, +, c↓↑, CV, AZ, b↓, A□, c↓↑, c↓, B÷, A□, AY, S, C↑, ↓, DV, AW, S, C↑, ↓, B÷, A÷, -, CX, EV, A√, a↑, d↓↑, X, ↓↑, C↓, -, A□, +, +, A□, V.

Code HK-2: Hewlett-Packard 9100B Codings for computing ΣX^2, ΣX, n, $\Sigma(X-\bar{X})^2$, s^2, s, $\bar{X}$, R and σ.

00 Clear, 01 0, 02 x→( ), 03 d, 04 Stop, 05 x→( ), 06 a, 07 x→( ), 08 c, 09 ↑, 0a x, 0b Acc+, 0c d, 0d ↑, 10 1, 11 +, 12 y→( ), 13 d, 14 RCL, 15 ↑, 16 d, 17 Stop, 18 IfFlag 19 3, 1a 4, 1b x→( ), 1c b, 1d ↑, 20 a, 21 If x=y, 22 0, 23 a, 24 If x<y, 25 y→( ), 26 a, 27 c, 28 If x=y, 29 0, 2a a, 2b If x>y, 2c y→( ), 2d c, 30 ↓, 31 Go To, 32 0, 33 9, 34 d, 35 ↑, 36 e, 37 X, 38 f, 39 ↑, 3a X, 3b ↓, 3c -, 3d d, 40 ÷, 41 y→( ), 42 b, 43 ÷ 44 y⇄( ), 45 b, 46 ↑, 47 1, 48 -, 49 ↓, 4a ÷, 4b ↓, 4c ↑, 4d √x, 50 PNT, 51 PNT, 52 f, 53 ↑, 54 d, 55 ÷, 56 a, 57 ↑, 58 c, 59 -, 5a b, 5b √x, 5c PNT, 5d PNT, 60 Go To, 61 0, 62 0, 63 End.

Code HK-3: Hewlett-Packard 9100B Codings for computing Mean Deviation about C, both Defining and Computing Formulas (C may be sample mean or median)

00 Clear, 01 0, 02 x→( ), 03 d, 04 x→( ), 05 c, 06 x→( ), 07 -, 08 f, 09 Stop, 0a x→( ), 0b b, 0c Stop, 0d If Flag, 10 4, 11 2, 12 x→( ), 13 a, 14 -, 15 f, 16 ↑, 17 1, 18 +, 19 y→( ), 1a -, 1b f, 1c a, 1d ↑, 20 b, 21 If x=y, 22 0, 23 c, 24 If x>y, 25 3, 26 4, 27 0, 28 Acc+, 29 c, 2a ↑, 2b 1, 2c +, 2d y→( ), 30 c, 31 Go To, 32 0, 33 c, 34 0, 35x⇄y, 36 Acc+, 37 d, 38 ↑, 39 1, 3a +, 3b y→( ), 3c d, 3d Go To, 40 0, 41 c, 42 d, 43 ↑, 44 c, 45 -, 46 b, 47 X, 48 ↑, 49 RCL, 4a -, 4b ↓, 4c x⇄y, 4d -, 50 x←( ), 51 -, 52 f, 53 ÷, 54 ↑, 55 b, 56 PNT, 57 PNT, 58 Clear, 59 Stop, 5a x→( ), 5b d, 5c Stop, 5d If Flag, 60 6, 61 b, 62 ↑, 63 d, 64 -, 65 |y|, 66 1, 67 Acc+, 68 Go To, 69 5, 6a c, 6b RCL, 6c ÷, 6d ↑, 70 d, 71 PNT, 72 PNT, 73 Go To, 74 0, 75 0, 76 End.

Code HK-12: Hewlett-Packard 9100B Codings for computing $\bar{\bar{X}}$, and $\bar{R}$, Control Limits for $\bar{X}$, and R-charts for Standard Known and Unknown.

00 Clear, 01 0, 02 x→( ), 03 d, 04 x→( ), 05 c, 06 3, 07 ↑, 08 Stop, 09 √x, 0a ÷, 0b y→( ), 0c b, 0d Stop, 10 x→( ), 11 a, 12 Stop, 13 x→( ), 14 -, 15 f, 16 Stop, 17 x→( ) 18 -, 19 e, 1a Stop, 1b x→( ), 1c -, 1d d, 20 Stop, 21 If Flag, 22 3, 23 8, 24 ↑, 25 Stop, 26 Acc+, 27 c 28 ↑, 29 1, 2a +, 2b y→( ), 2c c, 2d d, 30 ↑, 31 Stop, 32 +, 33 y→( ), 34 d, 35 Go To, 36 2, 37 0, 38 e, 39 ↑, 3a c, 3b ÷, 3c y→( ), 3d e, 40 f, 41 ↑, 42 c, 43 ÷, 44 y→( ), 45 f, 46 d, 47 ↑, 48 c, 49÷, 4a y→( ), 4b d, 4c e, 4d Roll ↓, 50 PNT, 51 Pnt, 52 f, 53 ↑, 54 a, 55 ÷, 56 b, 57 X, 58 e, 59 x⇄y 5a +, 5b ↑, 5c ↓, 5d -, 60 -, 61 c, 62 PNT, 63 PNT, 64 x←( ), 65 -, 66 f, 67 ↑, 68 a, 69 ÷, 6a 3, 6b X, 6c f, 6d X, 70 f, 71 x⇄y, 72+, 73 ↑, 74 ↓, 75 -, 76 -, 77 c, 78 PNT, 79 PNT, 7a If Flag, 7b 9, 7c a, 7d y⇄( ), 80 d, 81 y⇄( ), 82 f, 83 y⇄( ), 84 d, 85 ⇄( ), 86 a, 87 y⇄( ), 88 -, 89 e, 8a y⇄( ), 8b a, 8c y⇄( ), 8d -, 90 d, 91 y⇄( ), 92 -, 93 f, 94 y⇄( ), 95 -, 96 d, 97 Go To, 98 5, 99 2, 9a Go to, 9b -, 9c 0, 9d 0, -00 Stop, -01 x→( ), -02 e, -03 ↑, -04 Stop, -05 x→( ), -06 d, -07 ↑, -08 b, -09 X, -0a ↓, 0b x⇄y, -0c +, -0d ↑, -10 ↓, -11 -, -12-, -13 c, -14 PNT, -15 PNT, -16 a, -17 ↑, -18 d, -19 X, -1a x←( ), -1b -, -1c f, -1d X, -20 3, -21 X, -22 ↓, -23 +, -24 ↑, -25 ↓, -26 -, -27 -, -28 c, -29 PNT, -2a PNT, -2b If Flag, -2c +, 12d 0 -30 0, -31 y⇄( ), -32 -, -33 d, -34 y⇄( ), -35 -, -36 f, -37 y⇄( ), -38 -, -39 d, -3a y⇄( ), -3b a, -3c y⇄( ), -3d -, -40 e, -41 y⇄( ) -42 a, -43 Go To, -44 1, -45 6.

Code HK-14: Hewlett-Packard 9100B Codings for computing Control Limits for p and u-charts for Standard Known and Unknown.

00 Clear, 01 Stop, 02 If Flag, 03 0, 04 b, 05 ↑, 06 Stop, 07 Acc+, 08 Go To, 09 0, 0a 1 0b RCL, 0c x⇄y, 0d ÷, 10 ↑, 11 f, 12 Roll↑, 13 PNT, 14 PNT, 15 x→( ), 16 d, 17 Stop, 18 x→( ), 19 c, 12 ↑, 1b Stop, 1c x⇄y, 1d ÷, 20 y→( ), 21 b, 22 d, 23 ↑, 24 If Flag, 25 2, 26 b, 27 1, 28 x⇄y, 29 -, 2a X, 2b c, 2c ÷, 2d 9, 30 X, 31 d, 32 x⇄y, 33 √x, 34 +, 35 ↑, 36 ↓, 37 -, 38 -, 39 b, 3a PNT, 3b PNT, 3c Go To, 3d 0, 40 0, 41 End.

Code HK-15: Hewlett-Packard 9100B Codings for computing Control Limits for np and c-charts for Standard Known and Unknown.

00 Clear, 01 Stop, 02 x→( ), 03 d, 04 Stop, 05 If Flag, 06 1, 07 0, 08 ↑, 09 1, 0a Acc+, 0b Go To, 0c 0, 0d 4, 10 RCL, 11 ÷, 12 ↑, 13 ↓, 14 d, 15 ÷, 16 f, 17 Roll↑, 18 PNT, 19 PNT, 1a x→( ), 1b c, 1c x⇄y, 1d If Flag, 20 2, 21 9, 22 d, 23 ÷, 24 1, 25 x⇄y, 26 -, 27 c, 28 X, 29 9, 2a X, 2b c, 2c x⇄y, 2d √x, 30 + 31 ↑, 32 ↓, 33 -, 34 -, 35 f, 36 PNT 37 PNT, 38 Go To, 39 0, 3a 0, 3b End.

Code WK-2: Wang 700A Codings for computing $\bar{X}$, n, σ^2, σ, R, X_{min}, s^2, and s.

Mark, 0002, 0, ST DIR, 0001, St dir, 0002, St dir, 0003, Stop, St dir, 0007, St dir, 0008, Mark, 1508, + dir, 0001, X^2, + dir, 0002, 1, + dir, 0003, Re Y, 0001, Re dir, 0003 ÷, Stop, St dir, 0006, Re Y, 0008, -, Write A, Group II, St dir, 0008, Re Y, 0007, Skip Y<X, St dir, 0007, Search, 1508, Mark, 0, Re Y, 0003, Re dir, 0002, X, Re dir, 0001, X^2, -, St Y, 0002, Re Y, 0003, 1, -, Re dir, 0003, X, St Y, 0001, X^2, Re Y, 0002, ÷, ↓, √x, Stop, Re Y, 0008, Re dir, 0007, -, Stop, Re Y, 0002, Re dir, 0001, ÷, ↓, √x, Return,

Code WK-3: Wang 700A Codings for computing Mean Deviation about C, both Defining and Computing Formulas, (C may be sample mean or median).

Mark, 0003, 0, St dir, 0001, St dir 0002, St dir, 0003, St dir, 0004, Stop, St dir, 0005, Mark, 1509, Stop, Re Y, 0005, Skip Y<X, Search, 1510, + dir, 0003, 1, + dir, 0001, Re Y, 0001, Re dir, 0002, Search, 1509, Mark, 1510, + dir. 0004, 1, +dir, 0002, Re Y, 0001, Re dir, 0002, Search, 1509, Mark, 1, Re Y, 0002, Re dir, 0001, +, St Y, 0006, Re dir, 0001, - dir, 0002, Re dir, 0004, - dir, 0003, Re Y, 0002, Re dir, 0005, X, Re dir 0003, +, Re dir, 0006, ÷, Stop, Mark, 0103, 0, St dir, 0000, St dir, 0001, Stop, St dir 0002, Mark, 1511, Stop, ↑, Re dir, 0002, -, ↓, |X|, + dir, 0000, 1, +dir, 0001, Re Y, 0000, Re dir, 0001, ÷, Search, 1511,

Code WK-12: Wang 700 A Codings for computing $\bar{\sigma}$, k, $\bar{\bar{X}}$, $\bar{R}$ and Control Limits for $\bar{X}$, σ and R-charts for both Standard Known and Standard Unknown.

Mark, 0012, 0, St dir, 0000, St dir, 0001, St dir, 0002, St dir, 0003, Stop, √x, ↑, 3, ⇅, ÷, St Y, 0006, 0, ↑, Stop, St dir, 0007, Stop, St dir, 0008, Stop, St dir, 0009, Stop, St dir, 0010, Mark, 1512, Stop, + dir, 0000, Stop, + dir, 0001, Stop, + dir, 0002, 1, + dir, 0003, Search, 1512, Mark, 2, Re dir, 0003, ÷ dir, 0000, ÷ dir, 0001, ÷ dir, 0002, Re Y, 0000, Stop, Re Y, 0001, Re dir, 0002, Stop, Mark, 1513, Re Y, 0001, Redir, 0006, X, Re dir, 0007, ÷, Re dir, 0000, ⇅, +, St Y, 0011, -, -, ↓, Re Y, 0011, Stop, Re dir, 0008, St dir, 0005, Re dir, 0007, ÷ dir, 0005, 3, X dir, 0005, Re dir, 0001, X dir, 0005, Re Y, 0001, Re dir, 0005, +, St Y, 0012, -, -, ↓, Re Y, 0012, Stop, Re dir, 0001, ⇄ dir, 0002, ⇄ dir, 0001, Re dir, 0010, ⇄ dir, 0008, ⇄ dir, 0010, Re dir, 0009, ⇄ dir, 0007, ⇄ dir, 0009, Search, 1513, Mark, 3, 0, ↑, Stop, St dir, 0000, ↑, Stop, St dir, 0001, X dir, 0006, Re dir, 0006, Re Y, 0000, +, St Y, 0004, -, -, ↓, Re Y, 0004, Stop, Mark, 1514, Re dir, 0010, St dir, 0005, Re dir, 0001, X dir, 0005, 3, X dir, 0005, Re Y, 0009, Re dir, 0001, X, Re dir, 0005, +, St Y, 0012, -, -, ↓, Re Y, 0012, Stop, Re dir, 0008, ⇄ dir, 0010, ⇄ dir, 0008, Re dir, 0007, ⇄ dir, 0009, ⇄ dir, 0007, Search 1514,

Code WK-14: Wang 700A codings for computing Control Limits for p and u-charts for both Standard Known and Standard Unknown.

Mark, 0014, 0, St dir, 0000, St dir, 0001, Mark, 1515, Stop, + dir, 0000, Stop, + dir, 0001, Re dir, 0001, ↑, Re dir, 0000, Search, 1515, Mark, 4, ÷, ⇅, Stop, St dir, 0002, Mark, 1502, Stop, St dir, 0003, ↑, Stop, ⇅, ÷, Re dir, 0002, ⇅, Stop, 1, ⇅, -, X, Mark, 5, Re dir, 0003 ÷, 9, X, Re dir, 0002, ⇅, √x, -, St Y, 0004, +, Re dir, 0004, Search 1502,

Code WK-15: Wang 700A Codings for computing Control Limits for np and c-charts for both Standard Known and Standard Unknown.

Mark, 0015, 0, St dir, 0000, St dir, 0001, Stop, St dir, 0003, Mark, 1501, Stop, + dir, 0001, 1, + dir, 0000, Re dir, 0001, ↑, Re dir, 0000, Search, 1501, Mark, 6, ÷, Stop, Re dir, 0003, ÷, ⇅, Stop, Re Y, 0003, X, ⇅, St dir, 0002, Stop, 1, ⇅, -, Re dir, 0002, X, ↓, Mark, 7, √x, ↑, 3, X, re dir, 0002, ⇅, -, St Y, 0004, +, +, Re dir, 0004, Stop, End.

Table 1. November Mean Temperature, New York City (Central Park) for years 1900 to 1969 inclusive Cf. OK-1

	1900	49.1		1925	43.9		1950	48.4
	1901	39.7		1926	44.9		1951	43.5
(1)	1902	51.6	(6)	1927	49.2	(11)	1952	48.6
	1903	42.2		1928	47.4		1953	49.7
	1904	42.4		1929	46.2		1954	46.4
	1905	44.1		1930	45.5		1955	44.3
	1906	45.5		1931	51.9		1956	46.7
(2)	1907	46.2	(7)	1932	43.9	(12)	1957	49.4
	1908	46.8		1933	41.8		1958	47.9
	1909	49.5		1934	48.9		1959	45.8
	1910	42.3		1935	48.6		1960	49.7
	1911	42.7		1936	42.7		1961	48.8
(3)	1912	47.8	(8)	1937	46.4	(13)	1962	43.2
	1913	47.3		1938	48.3		1963	50.4
	1914	44.5		1939	43.7		1964	49.4
	1915	46.3		1940	45.3		1965	46.8
	1916	45.5		1941	50.0		1966	48.9
(4)	1917	41.6	(9)	1942	47.0	(14)	1967	42.5
	1918	46.6		1943	45.4		1968	46.9
	1919	45.2		1944	46.0		1969	46.4
	1920	44.4		1945	47.6			
	1921	44.7		1946	50.5			
(5)	1922	45.7	(10)	1947	44.2			
	1923	45.2		1948	52.4			
	1924	44.4		1949	46.3			

Data from U.S. Department of Commerce, Environmental Science Services Administration, Environmental Data Service, 30 Rockefeller Plaza, New York, New York 10020. [12]

For Code OK-1:

(Data coded by subtracting 40 from each x-values)

k	m'_k (for $u_i = X_i - 40$)	m_k	
1	6.3286	0	$\bar{X} = 6.3286 + 40 = 46.3286$
2	47.3049	7.2541	$\sigma = \sqrt{m_2} = 2.6933$
3	391.8126	0.6250	$a_3 = m_3/\sigma^3 = 0.0320$
4	3496.5948	133.5247	$a_4 = m_4/\sigma^4 = 2.5375$

For Code OK-4:

Table 3. Data from Table 1 Grouped in One-degree Intervals (Cf OK-4).

Class Interval	Class Marks, x_i	Coded Class Marks $u_i = x_i - 46.45$	Frequencies, f_i
3910 - 39.9	39.45	-7	1
40.0 - 40.9	40.45	-6	0
41.0 - 41.9	41.45	-5	2
42.0 - 42.9	42.45	-4	6
43.0 - 43.9	43.45	-3	5
44.0 - 44.9	44.45	-2	8
45.0 - 45.9	45.45	-1	9
46.0 - 46.9	46.45	0	13
47.0 - 47.9	47.45	1	6
48.0 - 48.9	48.45	2	7
49.0 - 49.9	49.45	3	7
50.0 - 50.9	50.45	4	3
51.0 - 51.9	51.45	5	2
52.0 - 52.9	52.45	6	1
		Total	70

Table 4. Data from Table 1 Grouped in Two Degree Intervals (Cf OK-4)

Class Intervals	Class Marks, x_i	Coded Class Marks $u_i = x_i - 45.95$	Frequencies, f_i
39.0 - 40.9	39.95	-6	1
41.0 - 42.9	41.95	-4	8
43.0 - 44.9	43.95	-2	13
45.0 - 46.9	45.95	0	22
47.0 - 48.9	47.95	2	13
49.0 - 50.9	49.95	4	10
51.0 - 52.9	51.95	6	3
		Total	70

For Code OK-11:

(a) $\bar{X}$-chart with standard known, $\bar{X}' \pm A\sigma'$.

(b) $\bar{X}$-chart with standard unknown, but σ' estimated by $\bar{\sigma}$, $\bar{\bar{X}} + 3\bar{\sigma}/c_2\sqrt{n} = \bar{\bar{X}} \pm A_1\bar{\sigma}$.

(c) $\bar{X}$-chart with standard unknown, but σ' estimated by $\bar{R}$, $\bar{\bar{X}} \pm 3\bar{R}/d_2\sqrt{n} = \bar{\bar{X}} \pm A_2\bar{R}$.

(d) $\bar{X}$-chart with standard unknown, but σ' estimated by s, $\bar{\bar{X}} \pm 3\ \bar{s}/c_4\sqrt{n} = \bar{\bar{X}} \pm A_3\ \bar{s}$.

(e) σ-chart with standard known, $(c_2 \pm 3c_3)\ \sigma' = (B_1,B_2)\ \sigma'$.

(f) σ-chart with standard unknown, $(1 \pm 3c_3/c_2)\ \bar{\sigma} = (B_3,B_4)\ \bar{\sigma}$.

(g) s-chart with standard known, $(c_4 \pm 3c_5)\ \sigma' = (B_5, B_6)\ \sigma'$.

(h) s-chart with standard unknown, $(1 \pm 3c_5/c_4)\ \bar{s} = (1 \pm 3c_3/c_2)\bar{s} = (B_3,B_4)\ \bar{s}$.

(i) R-chart with standard known, $(d_2 \pm 3d_3)\ \sigma' = (D_1,D_2)\sigma'$.

(j) R-chart with standard unknown, $(1 \pm 3d_3/d_2)R = (D_3,D_4)\bar{R}$.

For Code OK-13:

Table 8 Control Limits Using OK-12, basic constants

Ten Control Charts	Center Lines	Control Limits
(a) $\overline{X}$-chart, Standard known	$\overline{X}' = 46.3286$	$\overline{X} \pm 3\sigma'/\sqrt{n} = 42.716,\ 49.942 = \overline{x}' + A\sigma'$
(b) $\overline{X}$-chart, Standard Unknown		$\overline{X} \pm 3\overline{\sigma}/c_2\sqrt{n} = 42.677,\ 49.980 = \overline{\overline{X}} \pm A_1\overline{\sigma}$
(c) $\overline{X}$-chart, Standard Unknown	$\overline{X} = 46.3286$	$\overline{X} \pm 3\overline{R}/d_2\sqrt{n} = 42.695\text{-}49.963 = \overline{\overline{X}} \pm A_2\overline{R}$
(d) $\overline{X}$-chart, Standard Unknown		$\overline{X} \pm 3\overline{s}/c_4\sqrt{n} = 42.677,\ 49.980 = \overline{\overline{X}} \pm A_3\overline{s}$
(e) σ-chart, Standard known	$c_2\sigma' = 2.2641$	$(c_2 \pm 3c_3)\sigma' = 0, 4.730 = (B_1, B_2)\ \sigma'$
(f) σ-chart, Standard Unknown	$\overline{\sigma} = 2.2881$	$(1 \pm 3c_3/c_2)\overline{\sigma} = 0, 4.780 = (B_3, B_4)\overline{\sigma}$
(g) S-chart, Standard known	$c_4\sigma' = 2.5314$	$(c_4 \pm 3c_5)\sigma' = 0, 5.288 = (B_5, B_6)\sigma'$
(h) s-chart, Standard Unknown	$\overline{s} = 2.5581$	$(1 \pm 3c_5/c_4)\overline{s} = 0, 5.344 = (B_3, B_4)\ \overline{s}$
(i) R-chart, Standard known	$d_2\sigma' = 6.2637$	$(d_2 \pm 3d_3)\sigma' = 0, 13.245 = (D_1, D_2)\sigma'$
(j) R-chart, Standard Unknown	$\overline{R} = 6.3000$	$(1 \pm 3d_3/d_2)\ \overline{R} = 0, 13.321 = (D_3, D_4)\ \overline{R}$

Table 9a Some Corrections to ASTM Manual on Quality Control of Materials

Reference, location	Original Results	Corrected Results
Example 1, p. 79	$\overline{\overline{X}} \pm 3\overline{\sigma}/\sqrt{n} = 32.1, 35.9$ $(1 \pm 3/\sqrt{2n})\overline{\sigma} = 3.08, 5.72$	$\overline{\overline{X}} \pm 3\overline{\sigma}/c_2\sqrt{n} = \overline{\overline{X}} \pm A_1\overline{\sigma} = 32.105, 35.895$ $(1 \pm 3c_3/c_2)\overline{\sigma} = (B_3, B_4)\overline{\sigma} = 3.063, 5.737$
Example 3, p. 81 (Data coded by -0.5 and $X10^4$)	$\overline{\overline{X}} = -0.20,\ \overline{\sigma} = 2.3$ $\overline{\overline{X}} \pm A_1\overline{\sigma} = -3.4, 3.0$ $(B_3, B_4)\ \tilde{\sigma} = 0.1, 4.5$	$\overline{\overline{X}} = -0.18355,\ \overline{\sigma} = 2.7674$ $\overline{\overline{X}} \pm A_1\overline{\sigma} = -4.085+, 3.719$ $(B_3, B_4) = 0.084, 5.451$
Example 12, p. 90	$\overline{X}' \pm A\sigma' = 33.2,\ 36.8$ $(B_1, B_2)\sigma' = 2.94, 5.46$	$\overline{X}' \pm A\sigma' = 33.218,\ 36.782$ $(B_1, B_2)\sigma' = 2.880,\ 5.393$

Table 2. Statistics resulted frpm using Program OK-2 on Data in Table 1($u_i = x_i - 40$). Cf OK-2

Subgroup (years)	$\sum u$	$\sum u^2$	$n\sum u^2 - (\sum u)^2$	s^2	s	$\overline{x} = \overline{u} + 40$	σ	R
(1) 1900-4	25.0	228.06	515.30	25.765°	5.076	45.00	4.540	11.9
(2) 05-09	32.1	221.99	79.54	3.977	1.994	46.42	1.784	5.4
(3) 10-14	24.6	146.96	129.64	6.482	2.546	44.92	2.277	5.5
(4) 15-19	25.2	143.10	80.46	4.023	2.006	45.04	1.794	5.0
(5) 20-24	24.4	120.34	6.34	0.317	0.563	44.88	0.504	1.3
(6) 25-29	31.6	217.06	86.74	4.337	2.083	46.32	1.863	5.3
(7) 30-34	32.0	269.52	323.60	16.180	4.022	46.40	3.598	10.1
(8) 35-39	29.7	204.79	141.86	7.093	2.663	45.94	2.382	5.9
(9) 40-44	33.7	242.25°	75.56	3.778	1.944	46.74	1.739	4.7
(10) 45-49	41.0	379.10	214.50	10.725°	3.275-	48.20	2.929	8.2
(11) 50-54	36.6	291.82	119.54	5.977	2.445-	47.32	2.187	6.2
(12) 55-59	34.1	247.79	76.14	3.807	1.951	46.82	1.745+	5.1
(13) 60-64	41.5	378.29	169.20	8.460	2.909	48.30	2.602	7.2
(14) 65-69	31.5	220.27	109.10	5.455°	2.336	46.30	2.089	6.4
All Subgroups	443.0 (Total)	3,311.34 (Total)	35,544.80 (Total)	7.359 (Ave.)	2.713 (Ave.)	46.3286 (Ave.)	2.693 (Ave.)	12.7 (Ave.)

Table 5. For Normal Distribution, c_2 $E\sigma/\sigma'$, $c_3 = \sqrt{\mathrm{Var}\ \sigma/\sigma'}$ From OK-5)

n	c_2	c_3	$1/\sqrt{2n}$	n	c_2	c_3	$1/\sqrt{2n}$
2	0.564190	0.426251	0.500000°	26	0.970826	0.137968	0.138675+
3	0.723601	0.378243	0.408248	27	0.971918	0.135416	0.136083
4	0.797885-	0.336720	0.353553	28	0.972932	0.133001	0.133631
5	0.840749	0.305191	0.316228	29	0.973875-	0.130710	0.131306
6	0.868627	0.280750	0.288675+	30	0.974754	0.128534	0.129099
7	0.888203	0.261225-	0.267261	31	0.975576	0.126463	0.127000
8	0.902703	0.245208	0.250000°	32	0.976347	0.124489	0.125000°
9	0.913875-	0.231779	0.235702	33	0.977070	0.122604	0.123091
10	0.922746	0.220319	0.223607	34	0.977750-	0.120803	0.121286
11	0.929960	0.210394	0.213201	35	0.978391	0.119078	0.119523
12	0.935942	0.201692	0.204124	36	0.978996	0.117426	0.117851
13	0.940982	0.193982	0.196116	37	0.979568	0.115840	0.116248
14	0.945288	0.187090	0.188982	38	0.980110	0.114317	0.114708
15	0.949007	0.180882	0.182574	39	0.980624	0.112852	0.113228
16	0.952254	0.175251	0.176777	40	0.981112	0.111443	0.111803
17	0.955111	0.170114	0.171499	41	0.981576	0.110085-	0.110432
18	0.957646	0.165402	0.166666	42	0.982018	0.108775-	0.109109
19	0.959910	0.161061	0.162221	43	0.982439	0.107511	0.107833
20	0.961944	0.157045-	0.158114	44	0.982840	0.106289	0.106600
21	0.963782	0.153313	0.154303	45	0.983224	0.105109	0.105409
22	0.965451	0.149836	0.150756	46	0.983591	0.103967	0.104257
23	0.966972	0.146584	0.147442	47	0.983942	0.102862	0.103142
24	0.968365+	0.143536	0.144338	48	0.984279	0.101791	0.102062
25	0.969645+	0.140669	0.141421	49	0.984602	0.100753	0.101015 +
26	0.970826	0.137968	0.138675+	50	0.984912	0.099746	0.100000°

Table 6. $d_2 = ER/\sigma'$, $d_3 = \sqrt{\mathrm{Var}\ R/\sigma'}$, for Uniform, Exponential Distributions. Their d_2 -values and d_2 (for normal) are compared with d_2(max) and $\tilde{d}_2$ (max). Cf OK-7, 8, 9 and 10.

n	$d_{2,u}$	$d_{2,e}$	d_2	d_2(max)	$\tilde{d}_2$(max)	$d_{3,u}$	$d_{3,e}$
2	1.154701	1.000000	1.128379	1.154701	1.581139	0.816497	1.000000
3	1.732051	1.500000	1.692569	1.732051	1.870829	0.774597	1.118034
4	2.078461	1.833333	2.058751	2.083952	2.121320	0.692820	1.666667
5	2.309401	2.083333	2.325929	2.340126	2.345208	0.617213	1.193152
6	2.474358	2.283333	2.534413	2.553327	2.549510	0.553282	1.209798
7	2.598076	2.450000	2.704357	2.744140	2.738613	0.500000	1.221224
8	2.694301	2.592857	2.847201	2.920761	2.915476	0.455420	1.229552
9	2.771281	2.717857	2.970026	3.086854	3.082207	0.417786	1.235689
10	2.834265-	2.828968	3.077505+	3.244395+	3.240370	0.385695-	1.240874
11	2.886751	2.928968	3.172873	3.394664	3.391165-	0.358057	1.244897
12	2.931163	3.019877	3.258455+	3.538604	3.535534	0.334037	1.248212
13	2.969230	3.103211	3.335980	3.676955-	3.674235-	0.312984	1.250990
14	3.002221	3.180134	3.406763	3.810317	3.807887	0.294392	1.253353
15	3.031089	3.251562	3.471827	3.939193	3.937004	0.277859	1.255387
16	3.056560	3.318229	3.531983	4.064004	4.062019	0.263067	1.257156
17	3.079201	3.380729	3.587884	4.185111	4.183300	0.249756	1.258708
18	3.099459	3.439552	3.640064	4.302823	4.301163	0.237718	1.260082
19	3.117691	3.495108	3.688963	4.417410	4.415880	0.226779	1.261306
20	3.134187	3.547740	3.734950-	4.529108	4.527693	0.216796	1.262404
21	3.149183	3.597740	3.778336	4.638124	4.636809	0.207651	1.263394
22	3.162875+	3.645359	3.819384	4.744641	4.743416	0.199242	1,264291
23	3.175426	3.690813	3.858323	4.848826	4.847680	0.191485+	1.265107
24	3.186973	3.734291	3.895348	4.950822	4.949747	0.184307	1.265854
25	3.197632	3.775958	3.930629	5.050762	5.049752	0.177646	1.266540

(To be continued)

(Table 6 (Continued)

n	$d_{2,u}$	$d_{2,e}$	d_2	$d_2(max)$	$\tilde{d}_2(max)$	$d_{3,u}$	$d_{3,e}$
26	3.207501	3.815958	3.964315+	5.148767	5.147815+	0.171448	1.267171
27	3.216667	3.854420	3.996538	5.244943	5.244044	0.165667	1.267755-
28	3.225198	3.891457	4.027413	5.339390	5.338539	0.160261	1.268296
29	3.233162	3.927171	4.057044	5.432197	5.431390	0.155197	1.268799
30	3.240611	3.961654	4.085521	5.523447	5.522681	0.150442	1.269267
31	3.247595+	3.994987	4.112928	5.613216	5.612486	0.145969	1.269705
32	3.254156	4.027245+	4.139337	5.701573	5.700877	0.141753	1.270114
33	3.260331	4.058495+	4.164816	5.788583	5.787918	0.137774	1.270499
34	3.266153	4.088798	4.189425+	5.874304	5.873670	0.134012	1.270860
35	3.271652	4.118210	4.213218	5.958795-	5.958188	0.130449	1.271200
36	3.276853	4.146781	4.236246	6.042105+	6.041523	0.127071	1.271521
37	3.281780	4.174559	4.258553	6.124283	6.123724	0.123863	1.271825-
38	3.286455+	4.201586	4.280182	6.205373	6.204837	0.120813	1.272112
39	3.290897	4.227902	4.301171	6.285418	6.284903	0.117909	1.272384
40	3.295121	4.253543	4.321554	6.364457	6.363961	0.115141	1.272642
41	3.299144	4.278543	4.341364	6.442527	6.442049	0.112500-	1.272888
42	3.302981	4.302933	4.360631	6.519663	6.519202	0.109977	1.273122
43	3.306642	4.326743	4.379382	6.595899	6.595453	0.107565+	1.273344
44	3.310142	4.349998	4.397643	6.671262	6.670832	0.105256	1.273557
45	3.313489	4.372726	4.415438	6.745785-	6.745369	0.103045-	1.273759
46	3.316693	4.394948	4.432790	6.819493	6.819091	0.100924	1.273953
47	3.319764	4.416687	4.449717	6.892413	6.892024	0.098888	1.274139
48	3.322710	4.437964	4.466241	6.964572	6.964194	0.096833	1.274316
49	3.325538	4.458797	4.482378	7.035988	7.035624	0.095054	1.274487
50	3.328254	4.479205+	4.498146	7.106691	7.106335+	0.093246	1.274650-

Table 7a. Control-limit Factors for $\bar{x}$-chart Cf. OK-11

n	A	A_1	A_2	A_3	E_1	E_2	E_3
2	2.12132	3.75994	1.87997	2.65868	5.31736	2.65868	3.75994
3	1.73205+	2.39365+	1.02333	1.95441	4.14593	1.77245+	3.38514
4	1.50000	1.87997	0.72860	1.62810	3.75994	1.45719	3.25621
5	1.34164	1.59577	0.57682	1.42730	3.56825-	1.28981	3.19154
6	1.22474	1.40998	0.48325-	1.28713	3.45373	1.18371	3.15281
7	1.13389	1.27662	0.41928	1.18188	3.37751	1.10932	3.12696
8	1.06066	1.17498	0.37253	1.09910	3.32335+	1.05367	3.10871
9	1.00000°	1.09424	0.33670	1.03166	3.28273	1.01009	3.09498
10	0.94868	1.02811	0.30826	0.97535+	3.25117	0.97482	3.08433
11	0.09453	0.97266	0.28508	0.92739	3.22595-	0.94552	3.07582
12	0.86603	0.92530	0.26578	0.88591	3.20533	0.92068	3.06887
13	0.83205+	0.88424	0.24942	0.84955-	3.18816	0.89929	3.06308
14	0.80178	0.84819	0.23535+	0.81734	3.17364	0.88060	3.05819
15	0.77460	0.81622	0.22311	0.78854	3.16120	0.86410	3.05401
16	0.75000°	0.78761	0.21235-	0.76260	3.15042	0.84938	3.05038
17	0.72761	0.76180	0.20280	0.73906	3.14099	0.83615-	3.04721
18	0.70711	0.73838	0.19426	0.71758	3.13268	0.82416	3.04442
19	0.68825-	0.71699	0.18657	0.69787	3.12529	0.81324	3.04194
20	0.67082	0.69736	0.17961	0.67970	3.11868	0.80322	3.03972
21	0.65465+	0.67925+	0.17327	0.66288	3.11274	0.79400	3.03772
22	0.63960	0.66249	0.16746	0.64726	3.10736	0.78547	3.03591
23	0.62554	0.64691	0.16213	0.63269	3.10247	0.77754	3.03427
24	0.61237	0.63238	0.15721	0.61906	3.09801	0.77015	3.03278
25	0.60000°	0.61878	0.15265-	0.60628	3.09391	0.76324	3.03140

Table 7b. Control Limit Factors for σ-chart Cf. OK-11

n	c_2	c_3	B_1	B_2	B_3	B_4
2	0.564190	0.426251	0	1.84294	0	3.26653
3	0.723601	0.378243	0	1.85833	0	2.56817
4	0.797885+	0.336720	0	1.80804	0	2.26605-
5	0.840749	0.305191	0	1.75632	0	2.08900
6	0.868627	0.280751	0.02637	1.71088	0.03036	1.96964
7	0.888203	0.261225-	0.10453	1.67188	0.11768	1.88232
8	0.902703	0.245208	0.16708	1.63833	0.18509	1.81491
9	0.913875-	0.231779	0.21854	1.60921	0.23913	1.76087
10	0.922746	0.220319	0.26179	1.58370	0.28371	1.71629
11	0.929960	0.210394	0.29878	1.56114	0.32128	1.67872
12	0.935942	0.201692	0.33087	1.54102	0.35351	1.64649
13	0.940982	0.193982	0.35904	1.52293	0.38155+	1.61845-
14	0.945288	0.187090	0.38402	1.50656	0.40624	1.59376
15	0.949007	0.180882	0.40636	1.49165+	0.42820	1.57180
16	0.952254	0.175251	0.42650+	1.47801	0.44789	1.55211
17	0.955111	0.170114	0.44477	1.46545	0.46567	1.53433
18	0.957646	0.165402	0.46144	1.45385+	0.48185-	1.51815+
19	0.959910	0.161061	0.47673	1.44309	0.49664	1.50336
20	0.961944	0.157045-	0.49081	1.43308	0.51023	1.48977
21	0.963782	0.153313	0.50384	1.42372	0.52278	1.47722
22	0.965451	0.149836	0.51594	1.41496	0.53441	1.46559
23	0.966972	0.146584	0.52722	1.40672	0.54523	1.45471
24	0.968365+	0.143536	0.53776	1.39897	0.55533	1.44467
25	0.969645+	0.140669	0.54764	1.39165+	0.56478	1.43522

Table 7c. Control-limit Factors for s-Chart Cf. OK-11

n	c_4	c_5	B_5	B_6
2	0.797885-	0.602810	0	2.60632
3	0.886227	0.463251	0	2.27598
4	0.921318	0.388811	0	2.08775-
5	0.939986	0.341214	0	1.96363
6	0.951533	0.307547	0.02890	1.87417
7	0.959397	0.282155+	0.11293	1.80586
8	0.965030	0.262138	0.17862	1.75144
9	0.969311	0.245839	0.23179	1.70683
10	0.972659	0.232237	0.27595-	1.66937
11	0.975350-	0.220663	0.31336	1.63734
12	0.977559	0.210660	0.34558	1.60954
13	0.979406	0.201903	0.37370	1.58511
14	0.980971	0.194152	0.39851	1.56343
15	0.982316	0.187230	0.42063	1.54401
16	0.983483	0.180998	0.44049	1.52648
17	0.984506	0.175349	0.45846	1.51055+
18	0.985410	0.170198	0.47482	1.49600
19	0.986214	0.165475-	0.48979	1.48264
20	0.986934	0.161124	0.50356	1.47031
21	0.987583	0.157099	0.51628	1.45888
22	0.988170	0.153362	0.52808	1.44826
23	0.988704	0.149879	0.53907	1.43834
24	0.989192	0.146623	0.54932	1.42906
25	0.989640	0.143570	0.55893	1.42035-

Table 7d. Control-limit Factors for R-chart Cf. OK-11

n	d_2	d_3	D_1	D_2	D_3	D_4
2	1.128379	0.852503	0	3.68589	0	3.26653
3	1.692569	0.888369	0	4.35768	0	2.57459
4	2.058751	0.879810	0	4.69818	0	2.28205+
5	2.325929	0.864085-	0	4.91818	0	2.11450+
6	2.534413	0.848043	0	5.07854	0	2.00383
7	2.704357	0.833210	0.20473	5.20399	0.07570	1.92430
8	2.847201	0.819836	0.38769	5.30671	0.13617	1.86383
9	2.970026	0.807840	0.54651	5.39355-	0.18401	1.81599
10	3.077505+	0.797057	0.68633	5.46868	0.22302	1.77698
11	3.172873	0.787322	0.81091	5.53484	0.25557	1.74443
12	3.258455+	0.778486	0.92300	5.59391	0.28326	1.71674
13	3.335980	0.770425-	1.02470	5.64726	0.30717	1.69283
14	3.406763	0.763033	1.11766	5.69586	0.32807	1.67193
15	3.471827	0.756222	1.20316	5.74049	0.34655-	1.65345+
16	3.531983	0.749920	1.28222	5.78174	0.36303	1.63697
17	3.587884	0.744064	1.35569	5.82008	0.37785+	1.62215-
18	3.640064	0.738604	1.42425+	5.85588	0.39127	1.60873
19	3.688963	0.733457	1.48859	5.88933	0.40353	1.59647
20	3.734950-	0.728660	1.54897	5.92093	0.41472	1.58528
21	3.778336	0.724208	1.60571	5.95096	0.42498	1.57502
22	3.819384	0.719950+	1.65952	5.97924	0.43450+	1.56550-
23	3.858323	0.715924	1.71055+	6.00609	0.44334	1.55666
24	3.895348	0.712107	1.75903	6.03167	0.45157	1.54843
25	3.930629	0.708482	1.80518	6.05607	0.45926	1.54074

Table 9 ASTM [2, p. 85], Example 8 (p-chart) by OK-14

i	n_i	x_i	p_i(%)	$(LCL)_i$(%)	$(UCL)_i$(%)	i	n_i	x_i	p_i(%)	$(LCL)_i$(%)	$(UCL)_i$(%)
1	580	9	1.5517	0	2.8265-	17	330	2	0.6061	0	3.2991
2	550	7	1.2727	0	2.8653	18	640	4	0.6250°	0	2.7566
3	580	3	0.5172	0	2.8265-	19	580	7	1.2069	0	2.8265-
4	640	9	1.4062	0	2.7566	20	550	9	1.6364	0	2.8653
5	880	13	1.4773	0.1979	2.5532	21	510	7	1.3725+	0	2.9228
6	880	14	1.5909	0.1979	2.5532	22	640	12	1.8750°	0	2.7566
7	640	14	2.1875°	0	2.7566	23	200	7	3.5000°	0	3.8463
8	550	10	1.8182	0	2.8653	24	330	5	1.5152	0	3.2991
9	580	12	2.0690	0	2.8265-	25	880	18	2.0455-	0.1979	2.5532
10	880	14	1.5909	0.1979	2.5532	26	880	7	0.7955	0.1979	2.5532
11	800	6	0.7500°	0.1404	2.6107	27	800	8	1.0000	0.1404	2.6107
12	800	12	1.5000°	0.1404	2.6107	28	580	8	1.3793	0	2.8265-
13	580	7	1.2069	0	2.8265-	29	880	15	1.7045	0.1979	2.5532
14	580	11	1.8966	0	2.8265-	30	880	3	0.3409	0.1979	2.5532
15	550	5	0.9091	0	2.8653	31	330	5	1.5152	0	3.2991
16	330	4	1.2121	0	3.2991		19410	267	$\bar{p}$=1.3756		

Table 10. ASTM [2, p. 87] Example 10 (u-chart) by OK-14

i	n_i	c_i	u_i	$(LCL)_i$	$(UCL)_i$	i	n_i	c_i	u_i	$(LCL)_i$	$(UCL)_i$
(1)	20	(72)	(3.600)	1.2827	3.3173	11	25	47	1.880	1.3901	3.2099
2	20	38	1.900	1.2827	3.3173	12	25	55	2.200	1.3901	3.2099
3	40	76	1.900	1.5806	3.0194	13	25	49	1.960	1.3901	3.2099
4	25	35	1.400	1.3901	3.2099	14	25	62	2.480	1.3901	3.2099
5	25	62	2.480	1.3901	3.2099	15	25	71	2.840	1.3901	3.2099
(6)	25	(81)	(3.240)	1.3901	3.2099	16	20	47	2.350°	1.2827	3.3173
7	40	97	2.425°	1.5806	3.0194	17	20	41	2.050°	1.2827	3.3173
8	40	78	1.950°	1.5806	3.0194	18	20	52	2.600	1.2827	3.3173
9	40	103	2.575°	1.5806	3.0194	(19)	40	(128)	(3.200)	1.5806	3.0194
(10)	40	56	(1.400)	1.5806	3.0194	20	40	84	2.100	1.5806	3.0194
				Total			560	1,334	$\bar{u}$ = 2.300		

INDEX SERIAL NUMBER — 1102

PROBABILISTIC FATIGUE DESIGN ALTERNATIVE TO MINER'S CUMULATIVE DAMAGE RULE

Edward B. Haugen and Paul H. Wirsching

Department of Aerospace and Mechanical Engineering
The University of Arizona
Tucson, Arizona

Abstract

Classical methods, employing Miner's Cumulative Damage Rule, for the design of dynamic loaded mechanical elements are contrasted with probabilistic methods.

A probabilistic design procedure is developed (and applied) for the synthesis of mechanical components, with reliability R(N) specified, subject to narrow band random loading. Design problems involving fatigue due to constant amplitude sinusoidal loading are also analyzed.

Based on recent test results and the development in this paper, it can be postulated that broad band random loading provides a less severe operational environment than the same number of narrow band random loading peaks at comparable RMS value (RMS value is identical to the standard deviation where the mean value is zero), due to relatively fewer zero crossings. Thus, design requirements for narrow band random loading represent an upper bound on requirements for random loads environments.

NOTE: The development in this paper was in part accomplished under support by the U.S. Army Weapons Command, Small Arms Laboratory, Rock Island Arsenal.

Symbolism

S	allowable stress or strength, psi.
$\bar{S}$	sample mean strength, psi.
s_S	sample strength standard deviation, psi.
s	applied stress, psi.
$\bar{s}$	sample mean stress
s_s	sample stress standard deviation, psi.
RMS	s_s for zero mean value. Applies to narrow band random loading
R(N)	specified reliability after N cycles of loading
$(\bar{S}, s_S)$	description of material strength, psi
$(\bar{s}, s_s)$	description of static applied stress, psi.
$(\bar{s}, 0)$	description of constant amplitude dynamic applied stress, psi.
$(RMS, 0) = (s_s, 0)$	description of narrow band random applied stress, psi.
$(\bar{x}, s_x)$	mean value and standard deviation of the random variable X
f(S)	probability density function of strength
f(s)	probability density function of applied stress
ρ	correlation coefficient

Introduction

Often mechanical systems operate in dynamic load environments, wherein the phenomenon of fatigue may present a problem due to the requirements of efficient design vs. the avoidance of failure. In classical deterministic practice, design against fatigue has been heretofore highly dependent on linear cumulative damage concepts.

In this paper probabilistic methods will be presented to cope with dynamic random loading and possible fatigue effects in components, which avoid logical inconsistencies in classical practice.

Discussion

Miner's Rule

In the absence of other fatigue design theory, (in 1945), Miner* proposed empirical linear damage rules. A cumulative damage concept, they held that fatigue damage could be expressed in termsof the number of applied load cycles, n, divided by the number of cycles, N, to produce failure, at a given applied stress level, s, on the conventional constant amplitude S-N curve, see Figure 1(b).

In estimating fatigue damage due to narrow band random dynamic loading, a multi-valued loading record was idealized to a convenient number of force levels, associated with stress levels $s_1, s_2, \ldots, s_r$. The result was n_1 cycles at level s_1, n_2 cycles at level s_2, etc. The total number of load cycles experienced was $n = n_1 + n_2 + \ldots + n_r$. Failure of a component was postulated when the increments of damage summed to unity, i.e., when:[3]

$$\sum n/N = n_1/N_1 + n_2/N_2 + \ldots + n_r/N_r > 1. \quad (1)$$

Experience indicates that the linear cumulative damage rule [equation (1)] oversimplifies the fatigue failure phenomenon. The following observations suggested the need for a more realistic approach to mechanical design for dynamic loading and fatigue:[2]

1. Mechanical failures have been observed over a wide range of $\sum n/N$ values.
2. Fatigue strengths of engineering materials are random variables, also are non-linear functions of the number of load cycles (contrary to classical assumptions), see Figure 2.
3. Cycle life, N, for a given material and geometry, at a fixed stress level, displays wide variability (contrary to implications of the S-N curve).
4. Stress at failure (strength) corresponding to a specified cycle life, N, displays considerable variability (contrary to assumptions in classical theory).
5. With random dynamic loading, the order in which load intensities occur in a sequence can greatly influence cycles to failure [contrary to equation (1)].

* also Palmgren in 1924

Probabilistic Approach

In mechanical design for dynamic loads environments, the intent must be to avoid fatigue failure for a specified life with a specified probability (reliability). All design variables and associated phenomena must be treated as random variables, the variability in each becoming a design parameter.[1] The key to the development of rational design theory is found in the interpretation of materials behavior.

Even in a constant amplitude dynamic (zero mean) loads environment, a material displays widely varying cycle life from specimen to specimen over a series of tests.[3] A statistical description of material behavior requires a series of tests, at each of a number of constant amplitude stress values. From the data at each stress value, distributional parameters of cycles to failure are estimated. From samples of cycle life mean value and standard deviation estimators, best fit mean value and $\pm 3\sigma$ loci are plotted. The log-log plots of these loci provide a statistical S-N envelope. From such statistical S-N envelopes, the mean value and standard deviation of strength corresponding to any cycle life, N, may be estimated (Figure 2).

If Miner's theory provided a valid description of material behavior, then constant amplitude fatigue data, an abundance of which is available for many metallic materials, could be used to predict narrow band random fatigue. Unfortunately, because of the general failure of Miner's theory, separate fatigue tests for random loading must be performed. In a manner similar to that for constant amplitude dynamic loads, statistical S-N envelopes for material subject to narrow band random (zero mean) loading may be developed (Figure 3). For a material subject to narrow band random loading, the strength distributional characteristics may be estimated for a specific cycle life, N, from the narrow band statistical S-N envelope.

In a test series of narrow band random cycle life determinations, each specimen experiences a different random sequence of load intensities. Thus, the effect of loading history, including random order of stress amplitudes, is included in the statistics of the cycles to failure distributions at each RMS (standard deviation, σ) level.

The following postulations are employed in developing the design algorithm presented in this paper:

1. The statistical S-N envelope for each material, subjected to a specific type of dynamic loading, is unique.
2. For a specific cycle life, N, the mean value and standard deviation of strength can be estimated from the statistical S-N envelope.
3. The lower bound on mechanical component reliability is associated with normally distributed strength (for a specified cycle life, N) with a given mean value and standard deviation on strength, see Figure 4. NOTE: The likelihood of negative skewness is considered remote, based on observations.
4. For a specified cycle life, N, (see Figure 3), it is implied in the distribution of strength that the material has previously experienced and survived N-1 cycles of constant amplitude sinusoidal or narrow band random loading, whichever applies. NOTE: Thus, in design for a specified cycle life the strength distribution can be directly used without the need to examine effects of load cycles individually.

Having a priori distributional estimators for anticipated loading, materials behavior, geometry, and other relevant phenomena, the distributional parameters of the strength and applied stress functions are developed with required geometric size parameters as unknowns. The mathematical systems used for calculating the statistics of the functions[2] are the Algebra of Expectation and the Algebra of Normal Functions (see Appendix). With reliability, R(N), specified for surviving N loading cycles, the probabilistic criterion to be satisfied is equation (2):[1]

$$R(N) = P\{S > s\} = P\{S - s > 0\} \tag{2}$$

The reliability of a component is the probability that strength exceeds stress for the mission, equation (3):[1]

$$R(N) = P\{S - s > 0\} = \int_{-\infty}^{\infty} f(s)\left[\int_{-\infty}^{\infty} f(s)\, dS\right] ds. \tag{3}$$

For the case of normally distributed stress and strength, where $S - s = z$ is normally distributed

$$R(N) = P\{S - s > 0\} = P\{z > 0\} = \int_{0}^{\infty} f(z)\, dz\ , \tag{4}$$

which yields[2]

$$R(N) = \frac{1}{\sigma_z\sqrt{2\pi}} \int_{0}^{\infty} \exp\left[-\frac{(z-\mu_z)^2}{2\sigma_z^{\,2}}\right] dz \tag{5}$$

since z is normally distributed, to utilize standard normal tables, let

$$t = \frac{z - \mu_z}{\sigma_z}\ ; \qquad dz = \sigma_z\, dt\ .$$

Solving for new limits yields $(-\frac{\mu_z}{\sigma_z}$ and ∞).

Thus,

$$R(N) = \frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu_z}{\sigma_z}}^{\infty} e^{-t^2/2}\, dt = \int_{-\frac{\bar{S}-\bar{s}}{\sqrt{\sigma_S^2 - \sigma_s^2}}}^{\infty} e^{-t^2/2}\, dt \tag{6}$$

Equating lower limits

$$t = -\frac{\mu_z}{\sigma_z} = -\frac{\bar{S}-\bar{s}}{\sqrt{\sigma_S^{\,2} - \sigma_s^{\,2}}} \qquad (1,\ 2) \tag{7}$$

In postulation (3), it was implied that an assumption of normally distributed strength is usually conservative (would result in under estimations of component reliability). Further, the limiting form of a multiplicative series of n random variables is log normal, as n becomes large. It follows that, for a given mean value and standard deviation of stress, the log normal distributional form is almost always a conservative assumption, since applied stress is a

multiplicative function involving loads, geometry, and other random phenomena. The utilization of normal strength models and log normal stress models in design provides a reasonable and conservative model, since the exact distributional forms are usually not known.

The theory, methodology, and statistical algebra have been developed for design of mechanical components, when stress and strength are normally distributed.[1, 2] The combination of normal strength and log normal stress presents a problem that can be resolved by utilizing design curves[2] such as Figure 5. Assume a design situation where reliability is specified, R(N) = 0.9965. Enter Figure 5 along the line corresponding to 0.9965 and locate the point of intersection with the curve for log normal stress. If the reliability value directly above the normal stress (approximately R'(N) = 0.999) is used with normal theory, the resulting design will satisfy the original requirements.

Since any two parameter distribution is uniquely determined by two statistics, S and s are defined by mean value $\bar{S}$ and $\bar{s}$ and standard deviations σ_S and σ_s estimations. Thus, utilizing normal function algebra (Appendix) to express the parameters of S and s together with equation (7), unique design solutions can be obtained.

Design Procedure[2]

The following steps are required for the probabilistic design of mechanical components subject to dynamic use environments:

1. Model the relevant design variable and other engineering phenomena as random variables. The mean value and standard deviation estimators of each is required.
2. Utilize normal function algebra (see Appendix) or the algebra of expectation to calculate $\bar{S}$, s_S, $\bar{s}$, and s_s, each of which may be a function of more than one random variable. The actual distributions of the component random variables need not be known, since assumed normal strength (allowable stress) is a conservative limit and lognormal applied stress is a conservative limit.
3. For calculation purposes, modify the specified reliability R(N) associated with normal strength and lognormal stress, estimating a fictitious reliability R'(N) associated with normal strength and normal stress having the same parameters.
4. Design for R'(N) reliability using normal theory and methodology, which results in a component satisfying or exceeding the specified reliability R(N), see Figure 5.

Design Example

The design procedure will now be applied in an example.

Shown in Figure 6 is a model of a leaf spring to be manufactured from AISI 4340 alloy steel, whose mechanical behavioral properties (tensile ultimate strength, constant amplitude fatigue strength, and narrow band random fatigue strength) are given in Figure 3. In the examples, loading (whether static, constant amplitude sinusoidal, or narrow band random) is applied at the mid-point and perpendicular to the axis of symmetry of the spring. The finite life designs are for 100,000 cycles. In all cases, the specified probability of mission survival is R(N) = 0.9965.

The spring geometry is given in Figure 6. All dimensions are random variables, and the first number in each couple is the mean value estimator, the second being the standard deviation estimator.

The random variable dimension to be determined in each case is h. From known manufacturing tolerances the standard deviation on h is estimated as:

$$s_h \approx 0.015\,\bar{h}$$

In the cases of design for constant amplitude sinusoidal finite life and narrow band random finite life the frequency is specified as 1250 rpm.

Three cases of loading are considered in this example:

1. Static. $(\bar{P}, s_P) = (1{,}800;\ 100)$ lbs.
2. Constant Amplitude Sinusoidal.

$$(\bar{P}, s_P) = 1{,}800;\ 0) \text{ lbs.}$$

3. Narrow Band Random.

$$(\bar{P}, s_P) = (1237;\ 0) \text{ lbs.} = 0.707(1{,}800;\ 0).*$$

<u>Case 1.</u> Design the spring to sustain the static load $(\bar{P}, s_P) = (1{,}800;\ 100)$lbs. with a probability R = 0.9965.

Given: Spring geometry as shown in Figure 6. Allowable stress (strength) from Figure 3 is:

$$(\bar{S}, s_S) = (140{,}000;\ 5{,}600) \text{ psi.}$$

Solution: Since the spring cross-section will be prismatical and rectangular with $\ell >> \bar{b}$, the formula for extreme fiber stress is:

$$s = \frac{Mc}{I}$$

where

$$c = h/2 \quad \text{and } I = bh^3/12.$$

$$s = 6M/bh^2 \ (\text{lbs/in}^2)$$

First the mean value and standard deviation estimators $(\bar{M}, s_M)$ are calculated, by the formulas given in the appendix. In this case, M is the product of random variables P/2 and $\ell/2$. Since the estimators of a product of a random variable and a constant are

$$\overline{cx} = c\bar{x} \text{ and } s_{cx} = c\, s_x \ .$$

$$\frac{1}{2}(1{,}800;\ 100) = (900;\ 50) \text{ lbs.}$$

and

$$(\tfrac{1}{2})(\bar{\ell}, s_\ell) = (15.00;\ 0.0375) \text{ in.}$$

The moment, M, is,

$$(\bar{M}, s_M) = (900;\ 50)(15.00;\ 0.0375)$$

By the formulas for the moment estimators for a product:

$$(\bar{M}, s_M) = (13{,}500;\ 750.76) \text{ in. lbs.}$$

* For comparison, the RMS level of random loading was chosen the same as that of constant amplitude loading

The product 6M, of a random variable and a constant is

$$6(\bar{M}, s_M) = 6(13{,}500;\ 750.76)$$

$$= (81{,}00;\ 4504.56) \text{ in. lbs.}$$

Thus, letting

$$(\bar{u}, s_u) = \frac{(81{,}000;\ 4{,}504.56)}{(2.00;\ 0.035)}$$

By the formulas for the quotient of two random variables:

$$\bar{u} \approx \frac{81{,}000}{2.00} = 40{,}500$$

$$s_u \approx \sqrt{\frac{(81{,}000)^2(0.035)^2 + (4505.56)^2(2.00)^2}{(2.00)^4}}$$

$$= 2{,}220$$

The moment estimators for a random variable squared are

$$(\mu_{h^2}, \sigma_{h^2}) \approx (\bar{h}^2, 2\bar{h}\, s_h)$$

and

$$(\bar{s}, s_s) = \frac{(40{,}500;\ 2440)}{(\bar{h}^2, 2\bar{h}\, s_h)}$$

From processing considerations

$$s_h \approx 0.015\, \bar{h}$$

$$(\bar{s}, s_s) = \frac{(40{,}500;\ 2440)}{(\bar{h}^2, 0.03\bar{h}^2)} \text{ lbs./in}^2$$

Applying the formulas for the moment estimators of a quotient yields

$$(\bar{s}, s_s) = \left(\frac{40{,}500}{\bar{h}^2};\ \frac{2730}{\bar{h}^2}\right) \text{ lbs./in}^2$$

From Figure 5, for R(N) = 0.9965 associated with normal strength and lognormal stress, the problem is solved for $\bar{h}$ using normal theory with normal strength and stress and a reliability R'(N) = 0.999.

Corresponding to R'(N) = 0.999, from normal probability tables t ≈ -3.0. Substituting into equation (7),

$$-3.0 = -\frac{140{,}000 - \dfrac{40{,}500}{\bar{h}^2}}{\sqrt{(5600)^2 + \left(\dfrac{2730}{\bar{h}^2}\right)^2}}$$

This equation is solved for $\bar{h}$.

$$\bar{h} = .6020 \text{ in.}$$

$$s_h = 0.015\bar{h} = .0090 \text{ in.}$$

<u>Case 2</u>: Design the spring to sustain constant amplitude loading $(\bar{P}, s_P) = (1{,}800;\ 0)$ lbs. with a probability R(N) = 0.9965 for 100,000 cycles of loading.

Given: Spring geometry as shown in Figure 6. Allowable stress (Fatigue strength) at 100,000, from Figure 5:

$$(\bar{S}, s_S)_{100{,}000} = (85{,}880;\ 4{,}120) \text{ psi}$$

Solution: As in Case 1, the formula for extreme fiber stress is:

$$s = \frac{Mc}{I}, \quad \text{and} \quad s = 6M/bh^2 \ (\text{lbs./in}^2).$$

$$(\bar{M}, s_M) = \frac{(\bar{P}, s_P)}{2} \cdot \frac{(\bar{\ell}, s_\ell)}{2} \text{ in. lbs.}$$

$$(\bar{M}, s_M) = \frac{1}{4}(1{,}800;\ 0)(30.00,\ 0.075)$$

$$\bar{M} = \frac{1}{4}(1{,}800)(30.00) = 13{,}500 \text{ in.lbs.}$$

$$s_M = \frac{1}{4}\sqrt{(1{,}800)^2(0.075)^2 + (0)^2(30.00)^2 + (0)^2(0.075)^2}$$

$$= 33.75 \text{ in. lbs.}$$

Thus $(\bar{M}, s_M) = (13{,}500;\ 33.75)$ ln. lbs.

$6(\bar{N}, s_M) = (91{,}000;\ 202.5)$ in. lbs.

Let $(\bar{u}, s_u) = \dfrac{(81{,}000;\ 202.5)}{(2.00;\ 0.035)}$

$$\bar{u} \approx 40{,}500$$

$$s_u \approx \sqrt{\frac{(81{,}000)^2(0.035)^2 + (202.5)^2(2.00)^2}{(2.00)^4}} = 712.0$$

$$(\bar{u}, s_u) \approx (40{,}500;\ 712.0)$$

$$(\mu_{h^2}, s_{h^2}) \approx (\bar{h}^2, 2\bar{h}\, s_h)$$

$$(\bar{s}, s_s) = \frac{(40{,}500;\ 712)}{(\bar{h}^2, 0.03\bar{h}^2)} = \left(\frac{40{,}500}{\bar{h}^2};\frac{1410}{\bar{h}^2}\right) \text{ psi.}$$

Substituting into equation (7)

$$3.0 = \frac{85{,}880 - \dfrac{40{,}500}{\bar{h}^2}}{\sqrt{(4120)^2 + \left(\dfrac{1410}{\bar{h}^2}\right)^2}}.$$

Solving for $\bar{h}$

$$\bar{h} = .7520 \text{ in.}$$

$$s_h \approx 0.0113 \text{ in.}$$

NOTE: For the same conditions except $R(N) = 0.980$, $\bar{h} = 0.5250$ in., roughly equivalent to the result obtained for static loading ($\bar{h} = 0.597$ in.).

Case 3: Design the spring to sustain stationary gaussian narrow band random loading $(P, s_P) = (1273; 0)$ lbs., with a probability $R(N) = 0.9965$ for 10^5 cycles of loading. Here P refers to the RMS level of random loading. For purposes of comparison the RMS level of random loading in this case was chosen to be the same as the RMS level of the constant amplitude case. $s_P = 0$ is a necessary condition for the force to be stationary.

Given: Spring geometry as shown in Figure 6. Allowable RMS random fatigue strength at 10^5 cycles from Figure 3 (narrow band loading):

$$(\bar{S}, s_S)_{100,000} = (60{,}500;\ 2{,}670) \text{ psi.}$$

Solution: As in Cases 1 and 2, the formula for extreme fiver stress is

$$s = \frac{Mc}{I} \quad \text{and} \quad s = 6M/bh^2 \ (\text{lbs/in}^2).$$

$$(\bar{M}, s_M) = \frac{(\bar{P}, s_P)}{2} \cdot \frac{(\bar{\ell}, s_\ell)}{2} \text{ in. lbs.}$$

$$= \frac{1}{4}\,(1273;\ 0)(30.0;\ 0.075)$$

$$\bar{M} = \frac{1}{4}\,(1273)(30) = 9547 \text{ in. lbs.}$$

$$s_M = \frac{1}{4}\sqrt{(1273)^2(0.075)^2 + (0)^2(30)^2 + (0)^2(0.075)^2}$$

$$= 23.9 \text{ in. lbs.}$$

$$(\bar{M}, s_M) = (9547;\ 23.9)$$

$$6(\bar{M}, s_M) = (57{,}282;\ 143.4)$$

Let, $$(\bar{u}, s_u) = \frac{(57{,}282;\ 143.4)}{(2.00;\ 0.035)}$$

$$\bar{u} = \frac{57{,}282}{2} = 28{,}641$$

$$s_u = \sqrt{\frac{(57{,}282)^2(0.035)^2 + (143.4)^2(2.00)^2}{(2.00)^4}} = 502$$

$$(\bar{u}, s_u) = (28{,}641;\ 502)$$

$$(\bar{s}, s_s) = \frac{(28{,}641;\ 502)}{(\bar{h}^2,\ 0.03\bar{h}^2)} \text{ psi}$$

$$= \left(\frac{28641}{\bar{h}^2};\ \frac{990}{\bar{h}^2}\right) \text{ psi}$$

Substituting into equation (7):

$$3.0 = \frac{60500 - \left(\frac{28641}{\bar{h}^2}\right)}{\sqrt{(2670)^2 + \left(\frac{990}{\bar{h}^2}\right)^2}}\,.$$

Solving for $\bar{h}^2$ yields,

$$\bar{h}^2 = 0.572 \text{ in.}^2$$

And thus, $(\bar{h}, s_h) = (0.757;\ 0.011)$ in.

NOTE: The near equality of the mean values of $\bar{h}$ in cases 2 and 3 is accidental.

The fundamental natural frequency of the spring for case 3 is $(\bar{f}, s_f) = (16{,}360;\ 580)$ rpm. Thus, the likelihood of resonance is negligible.

Summary

The probabilistic design criterion presented in this paper includes the following unique considerations:

1. Loading is realistically modeled, including variability considerations.
2. Consideration of effects due to the order of random magnitudes of loading is included in the probabilistic method[2].
3. Materials behavior is realistically modeled as a statistical S-N envelope[2, 7].
4. Validity of the probabilistic design algorithm is not dependent on exact knowledge of the strength and stress distributional forms. Thus, pre-occupation with distributional form of material strength may be largely academic, the important design information being the mean value and standard deviation estimators[2].
5. The probabilistic design algorithm provides for designing for a random load environment to satisfy a stated measure of reliability, rather than simply analyzing an existing design.
6. It may be postulated that (see Figure 7) a narrow band operational environment is more severe than a comparable broad band random loads environment, for the same number of peaks and RMS value.[5]

Appendix

Summary of Binary Operations

Addition: $\mu_{x+y} = \mu_x + \mu_y$

$$\sigma_{x+y} = \sqrt{\sigma_x^2 + \sigma_y^2 + 2\rho\,\sigma_x\sigma_y}$$

Subtraction: $\mu_{x-y} = \mu_x - \mu_y$

$$\sigma_{x-y} = \sqrt{\sigma_x^2 + \sigma_y^2 + 2\rho\,\sigma_x\sigma_y}$$

Multiplication: $\mu_{xy} = \mu_x\mu_y + \rho\,\sigma_x\sigma_y$

$$\sigma_{xy} \approx [\mu_x^2\sigma_y^2 + \mu_y^2\sigma_x^2 + \sigma_x^2\sigma_y^2 + 2\rho\,\mu_x\mu_y\,\sigma_x\sigma_y + \rho^2\,\sigma_x^2\sigma_y^2]^{1/2}$$

Division: $\mu_{y/x} = \frac{\mu_y}{\mu_x} + \frac{\sigma_x \mu_y}{\mu_x^2}\left(\frac{\sigma_x}{\mu_x} - \rho\,\frac{\sigma_y}{\mu_y}\right)$

$$\sigma_{y/x} = \frac{\mu_y^2}{\mu_x^2}\left[\frac{\sigma_x^2}{\mu_x^2} + \frac{\sigma_y^2}{\mu_y^2} - 2\rho\,\frac{\sigma_x \sigma_y}{\mu_x \mu_y}\right]$$

References

1. Probabilistic Approaches to Design, E. B. Haugen, John Wiley & Sons Publishers, New York, 1968.
2. Probabilistic Mechanical Design, E.B. Haugen, The University of Arizona, Tucson, Arizona, 1972.
3. Stress, Strain and Strength, R.C. Juvinall, McGraw-Hill Book Co., New York, 1967.
4. Random Load Fatigue Testing: A State of the Art Survey, S.R. Swanson, Materials Research and Standards, Vol. 8, No. 4, ASTM, April 1968.
5. Fatigue of 2024-T3 Aluminum Alloy Due to Broad-Band and Narrow Band Random Loading, B.M. Hillberry, Iowa State University of Science & Tech., Ph.D. Dissertation, 1967.
6. The Fatigue of Aluminum Alloys Subject to Random Loading, R. Ikegami, University of California, Berkeley, Ph.D. Dissertation, 1969, Mechanical Engineering.
7. Statistical Metals Manual, E.B. Haugen and D. H. Mutti, The University of Arizona, Tucson, Arizona, 1972.

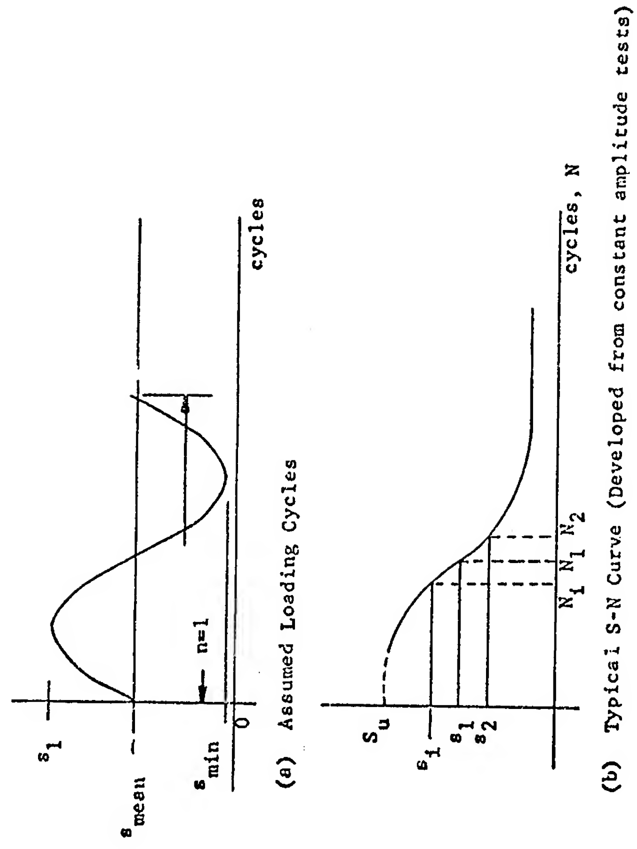


Figure 1.

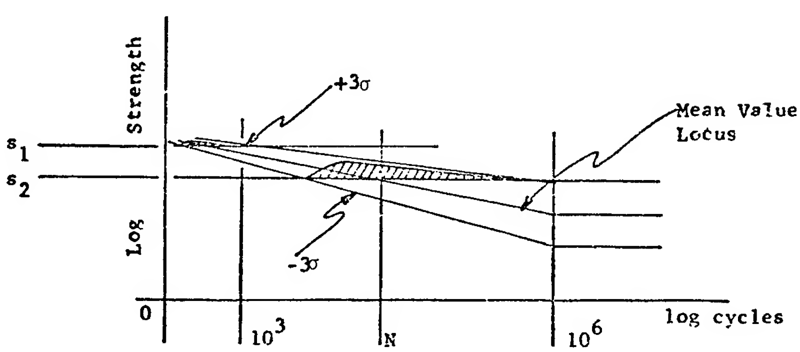


Figure 2. Statistical S-N Envelope

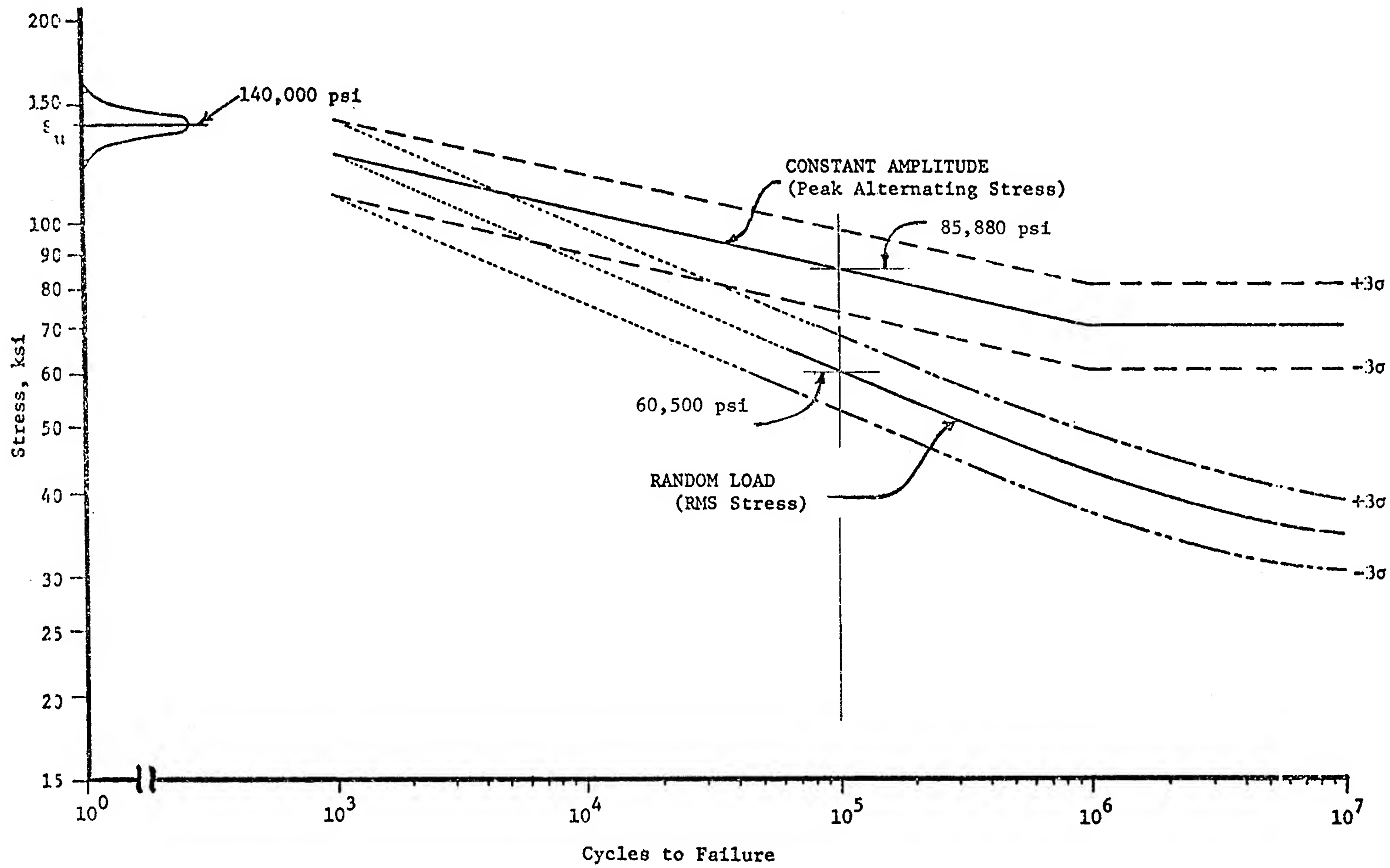


Figure 3. Statistical S-N Surface for SAE 4340 Steel Alloy (Derived from data from [4])

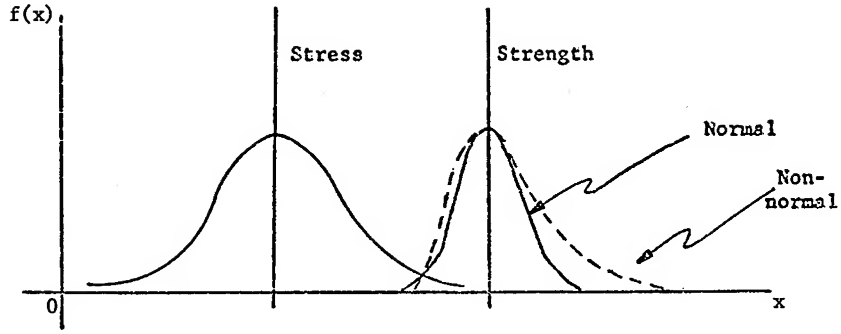
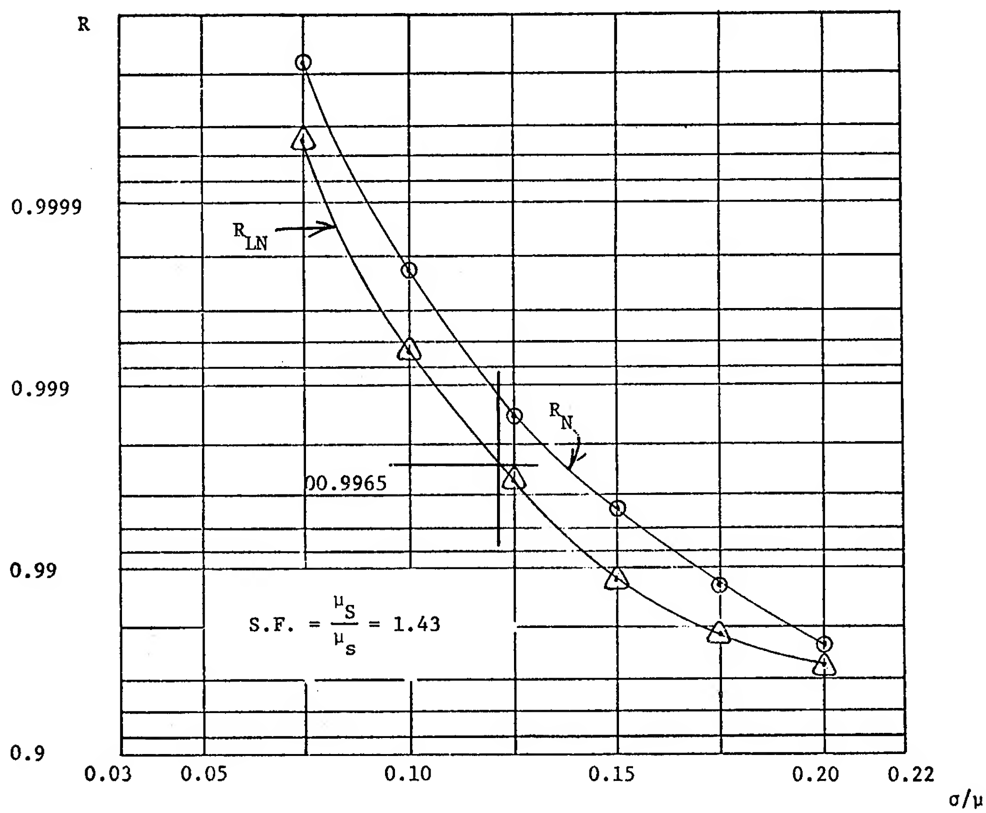


Figure 4. Stress-Strength Diagram [2]



Reliability versus Stress Coefficient of Variation
Normal Strength – (μ_S, σ_S) – (10.0000; 0.5000)
Normal and Log normal stress – μ_s = 7.0000 [2]

Figure 5

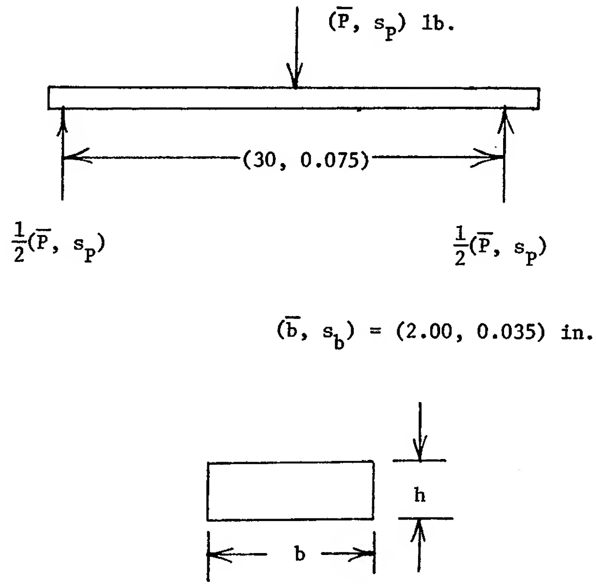


Figure 6 Leaf Spring [2]

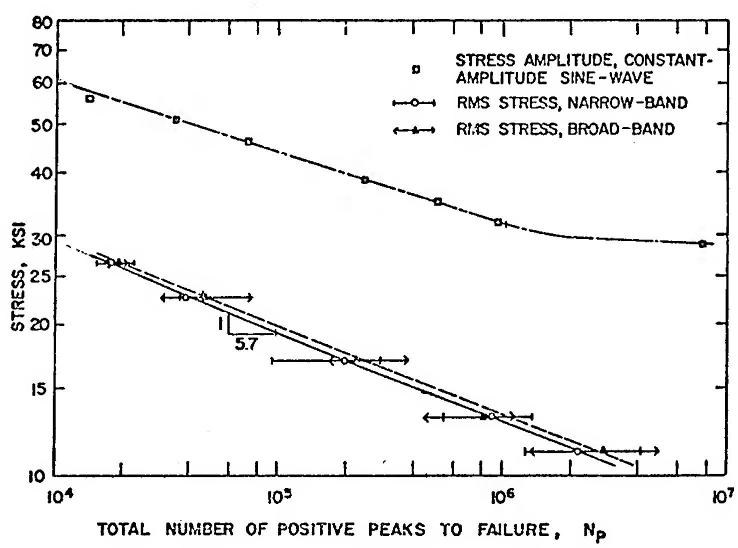


Figure 7. Random Fatigue Test Results [7]

A DESIGNER'S APPROACH TO THE FATIGUE FAILURE MECHANISM

INDEX SERIAL NUMBER — 1103

T. L. Salt
Aircraft Engine Group
General Electric Company
Lynn, Massachusetts

ABSTRACT

The assurance of structural reliability must be enhanced by a good understanding of the basic failure mechanism involved. Otherwise, the progress through to the provision of confident reliability guarantees may follow an expensive and tortuous path. A prime failure mechanism is fatigue and yet after many years of research and innumerable publications, this is still the subject of extensive investigation.

The current generally accepted engineering procedure is to confirm or adjust the design prediction for fatigue life as efficiently and as soon as possible by the utilization of component testing and the reduction of field experience. Indeed the efficient application of such a procedure has been the subject of previous papers including the application of Bayes Theorem.[1]

While it is certain that this type of testing will continue this paper reverts back to the basic understanding of the fatigue mechanism and provides an early description of a fresh approach and the interesting results obtained.

INTRODUCTION

The specification of structural reliability must include a supporting significant level of confidence and yet while fatigue is a serious failure mechanism it is generally accepted that the required level of confidence is difficult to obtain from design prediction. Indeed a recent paper by Grover[2] states that "The design to prevent fatigue failure is often speculative" and another by Duggan[3] says "the application of the data to design has still not reached a stage where the fatigue resistance of a component can be assessed with a high degree of confidence".

While there have been significant advances made by Manson[4] and Coffin[5] there is still little exception to the view that the designer still does not have a completely acceptable prediction method for general engineering components.

The pessimist might therefore conclude that, since to this time all the efforts of modern technology have been negative in producing an acceptable explanation of the fatigue mechanism, this fundamental behavior of materials must remain forever insoluble. To the contrary, the optimist can regard the present existence of so many pieces of evidence as a challenging opportunity to apply a fresh approach.

By accepting this challenge an approach has been derived which while still in initial development has provided several interesting results. While it presents a viewpoint different from that currently held by others, the magnitude of the problem and the possible significance of the results makes it opportune to bring it to the attention of those actively working in this field, even at this relatively early stage. Particularly as it already seems possible to generate a wide range of costly basic fatigue data from a relatively few measurements.

CORRELATIONS

An acceptable explanation for the fatigue mechanism and a mathematical application has to be consistent with the many known factors involved in the prediction of fatigue life, such as:

1. The distinction between strain-controlled and load-controlled specimen results even at high cyclic lives.

2. The effect of the alternating/mean stress ratio.

3. The derivation of a complete alternating/mean stress diagram in agreement with test data, obviating the need for the somewhat unsatisfactory empirical relationships of Goodman, Gerber and Soderberg.

4. The effective notch stress concentration K_f and its comparison with the theoretical stress concentration K_t in the evaluation of notch sensitivity.

5. The notch strengthening of some materials in ultimate tension.

6. The effect of grain size.

7. The effect of common material properties such as yield, ultimate, reduction of area and Young's modulus.

8. The effect of external environment such as temperature.

9. The limitations of the Miner/Palmgren/Langer cumulative damage rule.

10. The increase in life obtained by intermediate machining of parts having some fatigue damage.

11. The effect of component size.

12. The definition of endurance limit.

13. The non-propagating crack.

14. The possibility of obtaining a reasonable smooth bar strain cycling prediction from two or three basic parameters.

15. The dilemma of choosing strain or load cycling data as appropriate for a given design problem.

16. The different results from bending and axial-axial testing to the same surface stress.

17. The combination of the effects of external loads and internal thermal gradients.

18. The smaller statistical spread in failure results from notched specimens than for smooth specimens of the same material.

19. The definition of crack initiation.

20. The contention that significant cracks can be missed just before component failure.

INDUCTION

A first item selected is that of the comparative sensitivity of different materials to the influence of notches. This comparison may be described as the relationship between the apparent notch stress concentration (K_f) and the calculated theoretical notch concentration (K_t).

The value K_f is commonly defined as the ratio of load stress required to produce failure in a given number of cycles for a smooth specimen to the stress required to produce failure in the same number of cycles for its notched counterpart, Figure 1.

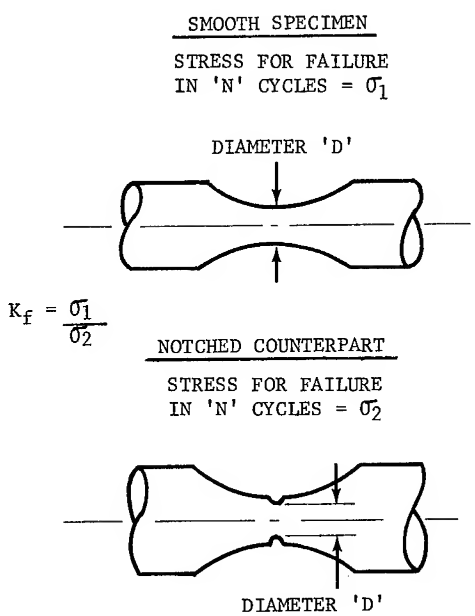
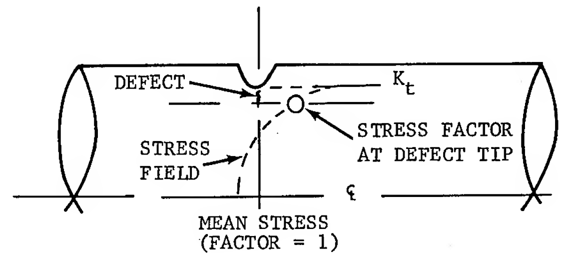


Figure 1 - Comparative Smooth & Notched Specimens

On the other hand the analytical value K_t is evaluated from a calculation of the peak value of stress (σ_p) at the notch surface compared with the average total section stress (σ_a) utilizing the principles of a theory of elasticity, i.e.: $K_t = \sigma_p/\sigma_a$.

It is well known that K_f is not equal to K_t and is generally believed to be always less than K_t. There is not even a constant relationship between K_f and K_t for a given material; for instance, there are differences arising from notch form, and a dependence upon the number of cycles considered, in the definition of K_f. Furthermore, there are materials such as cast irons which are so insensitive to processing notches that K_f will approach unity regardless of the values of K_t. The interpretation attributed to this behavior,[6] here, is that all common materials have inherent defects and those whose inherent defects are very large can hardly feel worse from the normal defects (notches) introduced in manufacture.

The consistency of this interpretation with notch behavior may be inferred by considering inherent defects extending from the surface of notches and relating their dimensions to the normal stress field situation created by the notch. The important criterion for progression of damage is the magnitude of the stress field at the extremity of the particular inherent defect. Figure 2.



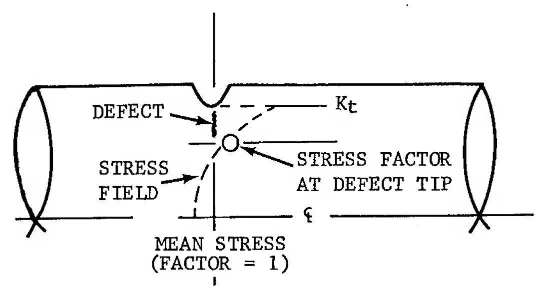


Figure 2 - Comparative Stress at Defect Tip. Materials with Short and Long Characteristic Defects (Notched Bars)

By definition the stress factor at the surface of the notch will be K_t, but by considering the presence of inherent defects, the critical position at the defect tip will feel a stress factor less than K_t. It may be deduced therefore that the size of the inherent defect is related to K_f and is in itself a measure of notch sensitivity. By the inference that all common materials contain inherent defects (except for whiskers of extreme purity), common materials will always demonstrate a K_f less than K_t. This inherent defect, probably better named as "a mean statistical characteristic defect", is a material property, and while having the usual statistical variations of other properties, will vary in a like manner with environment and processing history.

Secondly, it is important to discuss the discrepancies which occur in application of the Miner "rule"[7] as a means for assessing the cumulative damage due to the simultaneous application of loads of different levels. This rule, also attributed to Langer and Palmgren,[8 9] is most commonly seen as an equation and unfortunately the convenience of the equation as a design evaluation tool has often led to a misunderstanding of its basis. The equation states that failure occurs when:

$$\frac{n_1}{N_1} + \frac{n_2}{N_2} + \cdots\cdots + \frac{n_r}{N_r} = 1 \qquad (1)$$

where N_r is the number of cycles which would produce failure at loading level r and n_r is the actual number of cycles

applied at this level.

The discrepancies observed are generally related to an order-of-loading influence such that for two load levels there may be more or less life according to which load level is applied first. This observation would deny the prime inference of equation 1. that "damage is a function of the loading applied" and suggest that "damage is a function of the loading applied and the condition of the component when it is applied". This conclusion is consistent with the current recognition of fatigue as a multi-stage mechanism but this is hardly surprising when stress/strain relationships and creep relationships have been acknowledged for years to be of a stage nature. What is more important is that it implies that there is no single empirical formula which can completely describe the fatigue mechanism.

The proof of this statement may be deduced as follows:

The empirical formula implied would have the general form

$$D/_n = f(L) \tag{2}$$

where $D/_n$ is the damage per cycle and $f(L)$ would be some function of loading (exponential, polynomial or otherwise).

If the level of damage considered were failure (F) in N_1 cycles of load level L_1 then: -

$$F = N_1 \; f(L1) \tag{3}$$

or in general for failure in N_r cycles of load level L_r

$$F = N_r \; f(L_r) \tag{4}$$

for a composite of cycles $n_1, n_2 \ldots n_r$ of load levels $L_1, L_2 \ldots L_r$

$$F = n_1 \; f(L_1) + n_2 \; f(L_2) + \ldots + n_r \; f(L_r) \tag{5}$$

By substituting values of $f(L_r)$ from equation 4. in equation 5. gives

$$1 = \frac{n1}{N_1} + \frac{n2}{N_2} + \ldots\ldots + \frac{n_r}{N_r} \tag{6}$$

Equation 6 is obviously Miner's rule and is therefore a necessary condition for the validity of an algorithm of the type required.

This is sufficient at this stage of induction to establish a first fatigue mechanism model which would consider inherent defects extending to failure under the application of cyclic loads in a relationship with the characteristic of "stages". Such a relationship does, of course, exist from Fracture Mechanics investigations into the rate of crack propagation with the instantaneous stress intensity at the crack tip. A typical diagram of this type with some added descriptive nomenclature to indicate the stages is shown in Figure 3, stress intensity being a function of the stress field at the tip of the crack and the length of the crack.

It is necessary to make the assumption that inherent defects may be considered as inherent cracks even though they may be of the nature of grain boundaries or metallurgical precipitates and would not normally be categorized as cracks by the usual processing defect inspection. After making this assumption a diagram of the type shown in

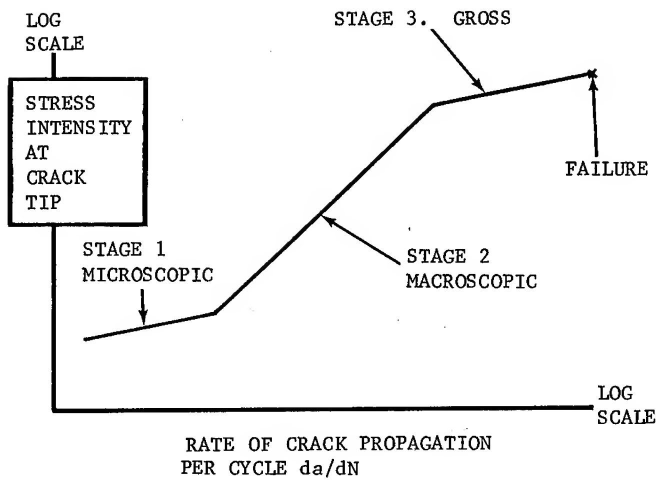


Figure 3. - Typical Fracture Mechanics Relationship Between Crack Growth and Stress Intensity

Figure 3 can be utilized in a discrete fashion by assuming progressively small increments of crack length increase and evaluating the incremental number of cycles which must have been consumed. The summation of all these cyclic increments up until the time that the stress intensity becomes so large that abrupt failure occurs will determine the cyclic life.

APPLICATION

While the procedure described is quite simple in concept and although there may be some nagging doubts relative to allocation of the same behavior to an inherent defect as to a crack, the successful application to test evidence should help to alleviate these doubts. There are, however, fundamental issues to be overcome before calculation can commence. The first is the determination of a mean statistical defect characteristic for the material under investigation and the second, which is a little more subtle, is to determine the appropriate stresses for the stress intensity calculations. A basic difficulty with this second issue is that fracture mechanics is "linear" implying difficulties in the consideration of plasticity; yet stresses extending into the plastic range cannot be avoided in fatigue analysis. Now is the time, therefore, to make an alternate interpretation of material behavior and yet maintain consistency with the previous induction made.

This step is quite simply to acknowledge that observed stress/strain curves are also the shape they are because of the material inherent defects, making it illogical to use stress/strain relationships which already have defects included and then introduce the defects again. After all, ultimate tensile ductility has been considered as a boundary condition of fatigue[5] and of course stress/strain curves show marked stages. It is necessary therefore to deduce what the stress/strain behavior of a material would be without its inherent defects before applying their known values in obtaining in one case the usual observed stress/strain curve and in another the usual observed fatigue characteristics. The effect of natural imperfections is deduced from the inferred behavior of the defect-free (perfect) material. "Perfection" is revealed by filaments of extreme purity (which have a linear stress/strain diagram), and even more fortunately can be deduced quite simply from a typical observed stress/strain diagram, Figure 4.

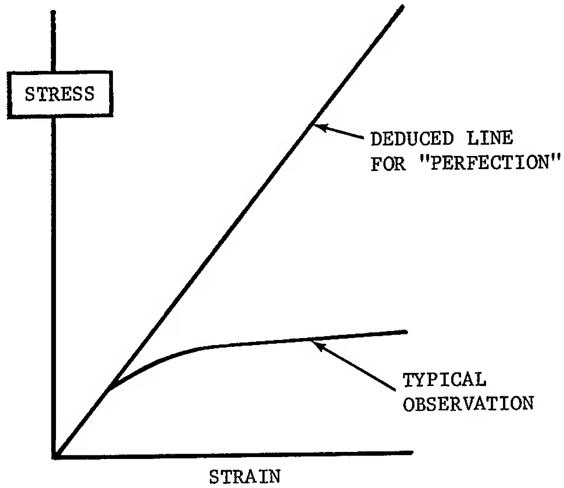


Figure 4. - Deduction of "Perfect" Material Line From Typical Observed Stress/Strain Diagram

The inference of this diagram is that for low stress levels there is generally a period where the natural defects are virtually ineffective and "elasticity" may be assumed (Hooke's Law) but as the stress increases their influence becomes more marked, giving the usual observed behavior. The initial slope, however, (Young's Modulus) provides a close approximation to the required diagram. In all calculation of stress intensity, therefore, the equivalent linear diagram representative of perfection can be used and effectively dispense with any concern relative to plasticity. This assumption is consistent with the behavior of wood, since it is known that knots do not influence stiffness or the elastic limit of beams. However, the range of stress between the elastic limit and modulus of rupture is seriously influenced by knots, crooked grain and other defects.

DETERMINATION OF MEAN STATISTICAL DEFECT

While there is obviously a wide range of experimental evidence which might be used to determine the mean statistical defect characteristic of a material in its processed form and operating environment, the arbitrary choice made was a low cycle fatigue result (1300 cycles to failure for a super alloy steel, smooth test specimen, cycled between large values of positive and negative strain). By using the fracture mechanics diagram for this material (Figure 3), it was possible to try various sizes of the defect as a starting length "a" until 1300 cycles to failure was obtained. Since this initial material was of high strength, the value obtained for the characteristic was the comparatively short length of .003 inches and for convenience of discussion was given the number of 30 where number = defect length in inches X 10000. Having the number established it was a simple matter to select other strain boundaries and generate a complete cycles/strain diagram. The total diagram obtained was an excellent correlation with observation, Figure 5.

Figure 5 was obtained for the particular conditions of fully reversed strain otherwise described as an "A" ratio of infinity where A ratio = alternating strain/mean strain. It is obviously essential to be able to consider any other A ratio and to make this possible, the assumption was introduced that stress intensity varies in general as a function of $(\text{max stress})^n \times (\text{stress range})^m$ where $n = 2$ and $m = 2\frac{1}{2}$. There is a correlation of this assumption to the old 4th power law of crack propagation when max stress and range are equal[10], and to the values of 2 and 2 which may be deduced from a publication of Walker[11]. Further investigation into the values of these indices is probably required, particularly the choice of $2\frac{1}{2}$ rather than 2 for m, but at present the values chosen seem to be the best. Introducing this refinement gave good correlation with A ratio of unity information.

Before moving to the alternative of load controlled testing, it is perhaps useful to provide a diagram to illustrate the derivation of the "perfection" stress/strain line under strain cycling conditions. For this purpose an A ratio of unity has been chosen for an assumed simple "elastic/plastic" material. The required diagram is shown in Figure 6.

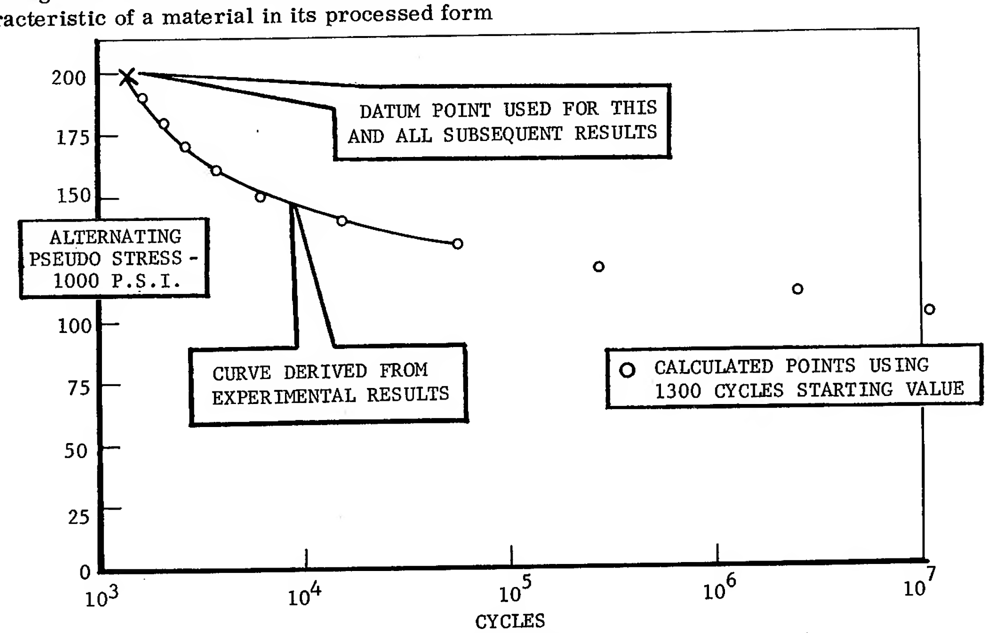


Figure 5. - Comparison of Calculated Points with Experimentally Generated Curve Superalloy - Strain Cycling (Pseudo Stress = Strain X E)

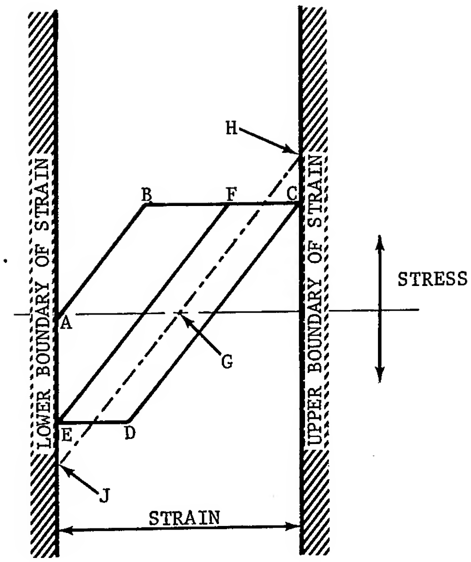


Figure 6. - Derivation of "Perfect" Stress/Strain Line, Strain Cycling A = 1

Figure 6 shows the sequence of cycling between the strain controlled barriers through A, B, C, D, E, F and back to C such that the hysteresis loop EFCD is established, whence the required "perfection" stress/strain relationship is derived as the line JGH. In practice an observed stress/strain relationship is more complex (3 stages ?) and requires several cycles to establish the needed line.

LOAD CONTROL

While strain-controlled test data is common, there are occasions when load-controlled data is required. The logical next step is therefore to calculate this type of data using the inherent defect characteristic value previously derived. Other than producing an analogous diagram to Figure 6 for boundaries of stress rather than strain, a small adjustment is necessary to account for the increase in stress with damage which occurs in this alternative system. It follows that, as deteriorated material is lost, the stress will increase for constant load but, with strain control, the load reduces proportionally with the area reduction Figure 7.

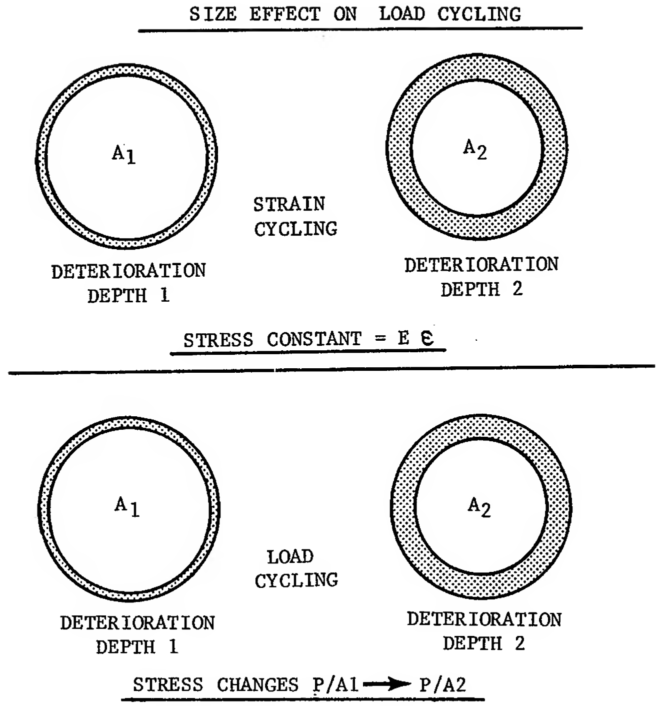


Figure 7. - Size Effect on Load Cycling

While the calculations for individual "A" ratios were again in very good agreement with test, it was considered more important to show a calculated mean stress/alternating stress diagram and compare it with the traditional empirical suggestions of Goodman/Gerber/Soderberg and with experimentally derived diagrams. This comparison is shown in Figure 8 and is consistent with the shape seen from the testing of many materials.

The approach described was repeated for a second super-alloy steel but of lower strength than the first. The characteristic defect size at the temperature considered was found to be 0.0092 (A number of 92). Agreement was obtained with both strain and load cycling data as before and as a further exercise a total mean stress/alternating stress diagram was calculated for this material, Figure 9.

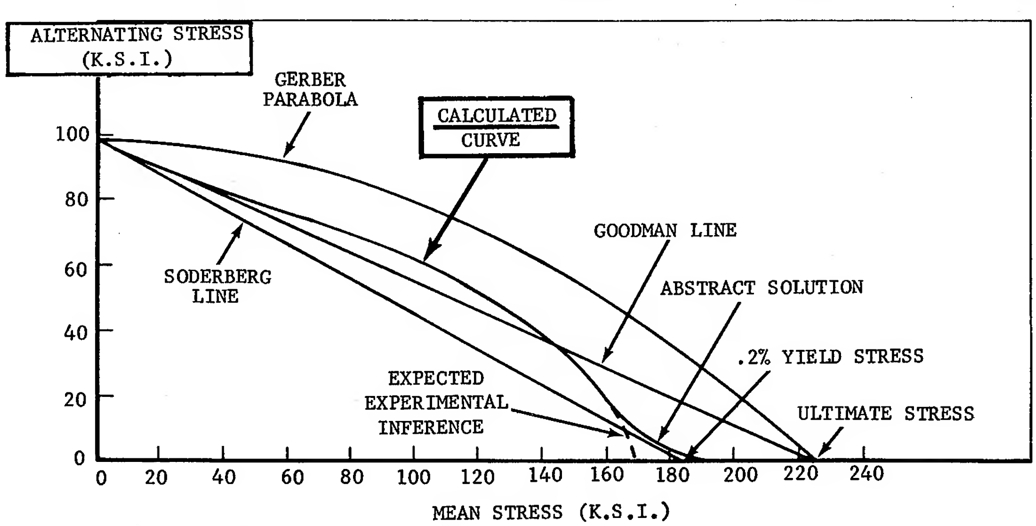


Figure 8. - Calculated Mean Stress/Alternating Stress Limits

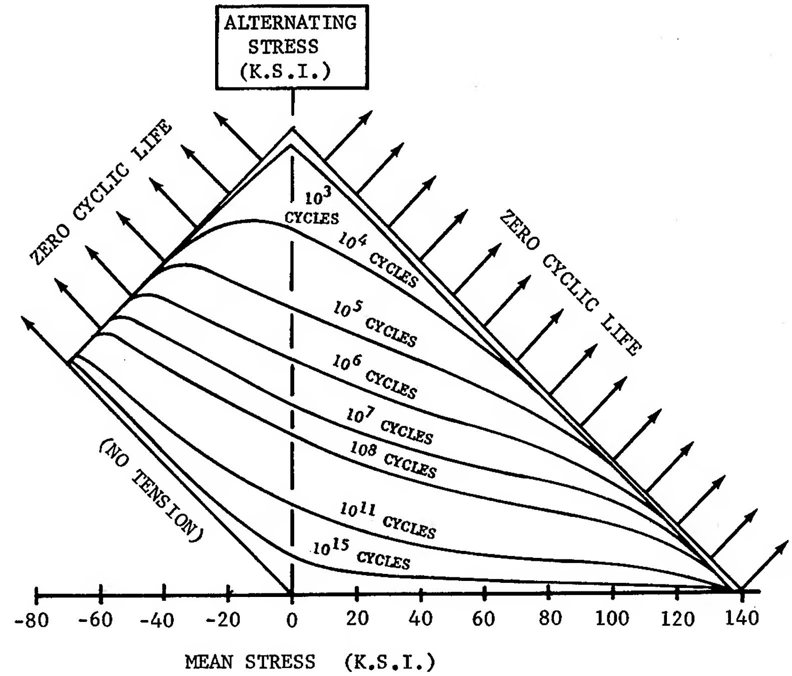


Figure 9. - Smooth Specimen Mean Stress/Alternating Stress Diagram Limits for Failure Due to Cyclic Loading

Beyond providing a general agreement with a few data points available in the negative mean stress regime, it has the beauty of showing the family of curves which must exist within the boundaries of zero and very long life. Indeed, the asymmetry of these boundaries explains the inflexions observed (mostly in the positive quadrant) in individual curves as a necessary requirement for them to fit the pattern.

NOTCHED BARS

Since the approach described has until this time been applied only to smooth bars, it is extremely important to move to notched bars. This is a most crucial issue as design components predominantly fail at notched locations and it is there where the critical predictions are made. No new material test data was introduced but it was necessary to recognize that a notched bar under load cycling experiences periods of either strain or load cycling, and to decide when the change occurs. The significance of this is extremely important as a designer traditionally may have basic data from only the individual alternatives and yet there may be an infinite selection of pieces of both. The application of the rule derived is as follows:

a) Determine the stress field in the vicinity of the notch by "elastic" stress analysis as factors of the mean stress which would be calculated for the section without regard for the stress concentration. Express these stress field factors as a function of depth from the surface such that at any given depth X the function will have the value F1.

$$\text{Stress}_x = \text{Mean stress} \times F_1 \qquad (7)$$

b) Determine the mean stress variation with deterioration as discussed previously, Figure 7. This relationship may be expressed as:

$$\text{Stress}_x = \text{Mean stress} \times F2 \qquad (8)$$

c) Load cycling will commence when F2 becomes equal to or greater than F1.

Applying this rule produced a relationship between K_f and number of cycles for three materials considered. All were of the same shape and consistent with the shape observed for many alloys[12]. One shape obtained and that of an aluminum alloy given in this reference are compared in Figure 10, where it must be noted that the dotted line, not the solid line, is the quoted general trend. In addition, it was shown that both of the super-alloys considered would have significant notch strengthening at very low cycles ($K_f < 1.0$) which indeed is the observed case.

STATUS

Current investigations have suggested that the approach described can be made more efficient by a generalization of the fracture mechanics diagram of the type shown in Figure 3 into a common parametric form which would be deduced by specification of only two more simple fatigue test values. The success of these investigations would indicate a considerable simplification in the method as a general practice.

The design approach to fatigue prediction would no longer involve individual reference to test bar data which can never be fully inclusive, but for each critical area would involve a prediction from the appropriate stress analysis and knowledge of the mean statistical defect characteristic of the material in the appropriate environment. Design analysis, on an experimental basis, of this type, has been initiated and is already extremely promising, some of the results obtained being intended for the subject of a later paper. Re-assessment of the listed "CORRELATIONS" is continuing and so far they have been found to be consistent with the approach.

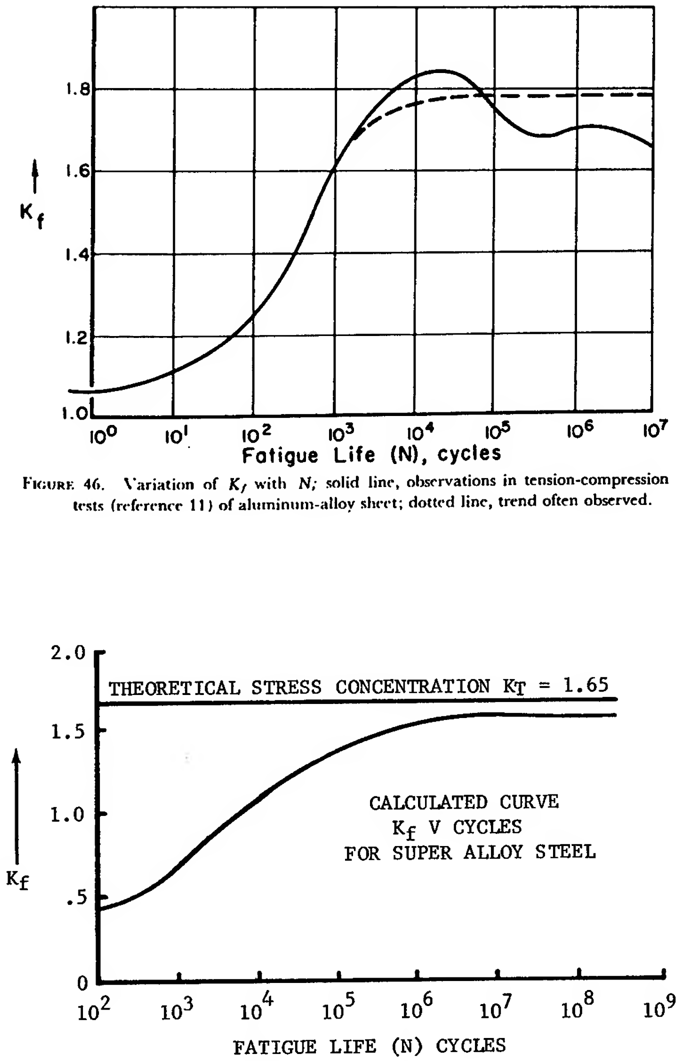


Figure 10. - Comparison Between Observed Variation of K_f with Cycles (Aluminum) and Calculation for Super-Alloy Steel

OVERALL REVIEW

It is very important to review why the approach described appears successful when in some respects it seems inconsistent with some current views, particularly since it has been done from what is commonly believed to be only a "crack propagation" model with no regard for a prior "crack initiation" model.

While crack initiation has been the crux of much argument in definition, it cannot be denied that surface micro-cracks are observed very early in the fatigue process and test observations have shown a sudden change in response easily related to a theory of micro-crack amalgamation into a definite advancing crack front. Yet the results obtained indicate that in failure calculations there is no need for an allocation of additional cycles for any such a preliminary mechanism. Furthermore by the approach described a strip of material of average thickness equal to the dimension of its inherent defect characteristic would fail very quickly with application of cycles and this is clearly not the case. Also, single crystals do fail. On the other side, ultimate tension which has been regarded as fatigue in a 1/4 cycle is accepted as a process of internal failure prior to surface cleavage, and there are examples of surface treatment causing internal "crack initiation".

The correlation of these apparently inconsistent facts can, however, be rationalized by an hypothesis that there are two mechanisms acting at the same time, such that in general components and test specimens the cycles utilized in the crack initiation mechanism are the same cycles used in the defect extension mechanism and hence do not have to be counted in failure prediction. The logic of this hypothesis is illustrated in Figure 11 and extended to the suggested situation for a general component where defect extension is predominant, to ultimate tension where slippage is predominant in failure prediction, and to single crystals where defect extension does not exist.

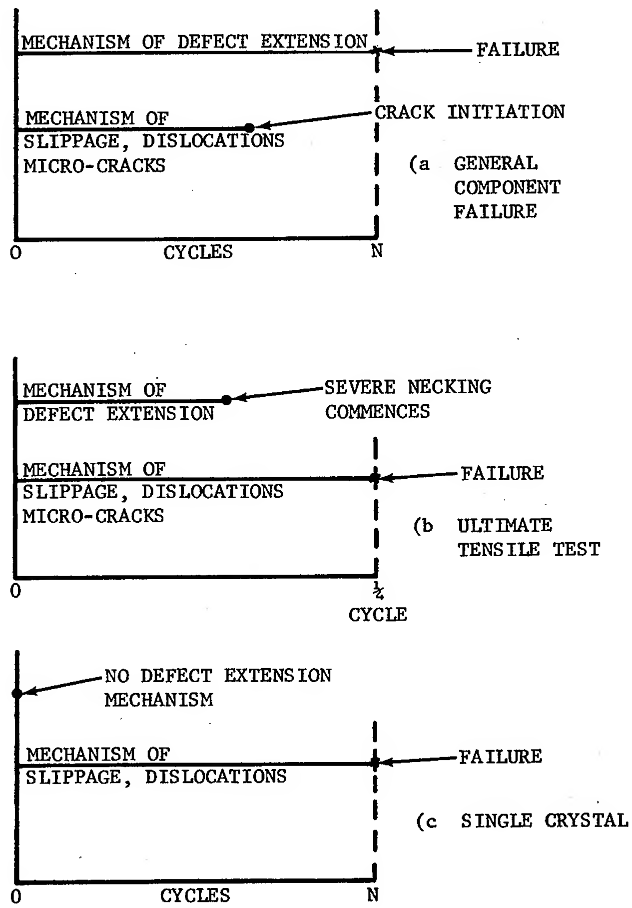


Figure 11. - Hypothesis of Coincident Mechanisms

Further validity for this hypothesis may be inferred from observations of high striation density in fatigue failures to a depth not inconsistent with the magnitude of an inferred characteristic defect, by the sudden discovery of a crack of extreme length beyond the characteristic defect size and by the conclusion that the inherent defect will be very large at high temperatures making the slippage mechanism predominant in creep. This deduction relative to high temperature could introduce an alternate approach to the consideration of the degradation by "hold time" but this still has to be confirmed.

CONCLUSIONS

1. A fresh approach to the understanding of the fatigue failure mechanism has been described which has already produced interesting new results such as the calculation of a total mean stress/alternating stress diagram and the effective notch concentration factor K_f.

2. The derivations obtained have been checked against twenty significant factors related to fatigue, and so far have been found to be quite consistent.

3. It is important to evaluate this approach in the context of specific component designs. Only this can determine its true value in the prediction of structural reliability.

REFERENCES

1. Salt, T. L., "A Routine Bayesian Assessment and Modification of Prediction", SAE/AIAA/ASME Ninth Annual Conference of Reliability and Maintainability, July 1970.

2. Grover, H., "Using Fatigue Information in Design", SESA Fall Meeting, October 1971, and published in the Journal of Experimental Mechanics, May 1972.

3. Duggan, T. V., "Current Trends in Fatigue Research", The Chartered Mechanical Engineer, November 1970.

4. Manson, S. S., "Fatigue: A Complex Subject - Some Simple Approximations", 1965 Experimental Mechanics SESA, Vol. 5, p. 193.

5. Coffin, L. F. and Tavernelli, J.D., "Cyclic Straining and Fatigue of Metals", Transactions of AIME, October 1959.

6. Orowan, E., "Proceedings of the Royal Society", 1939.

7. Miner, M.A., "Cumulative Damage in Fatigue", Journal of Applied Mechanics, 1945.

8. Langer, B. F., "Design of Pressure Vessels for Low Cycle Fatigue", Journal of Basic Engineering (Transactions-ASME), 1962.

9. Palmgren, A., "Life of Ball Bearings", 1924.

10. Johnson, H. H. and Paris, P. C., "Sub-Critical Flaw Growth", National Symposium on Fracture Mechanics, Lehigh University, June 1967.

11. Walker, E.K., "The Effect of Stress Ratio During Crack Propagation and Fatigue for 2024-T3 and 7075-T6 Aluminum", ASTM Winter Meeting, 1969.

12. Several Authors, "The Fatigue of Metals and Structures", NavWeps 00-25-534.

ACKNOWLEDGEMENTS

To Mrs. Olive King for such excellent help in computational aspects of this work.

CYCLES-TO-FAILURE AND STRESS-TO-FAILURE WEIBULL DISTRIBUTIONS IN STEEL WIRE FATIGUE

Guy Robert Wagner
Bell Telephone Laboratories
Denver, Colorado 80221

Charles R. Mischke
Department of Mechanical Engineering
and Engineering Research Institute
Iowa State University
Ames, Iowa 50010

Summary

Based on 520 specimens of music wire, a three-parameter Weibull statistical model is fitted to the cycles-to-failure distribution and the stress-to-failure distribution of a conventional SN-curve. The shape of the distribution is not constant over the experimental range (b varies). The cycles-to-failure distribution exhibits both a positive and a negative skewness in the experimental range, illuminating the reasons for some differences experienced in fitting Gaussian and log-normal models to this distribution. The stress-to-failure distribution exhibits only negative skewness in the experimental range.

Symbols and Notation

Symbol	Description
b	Weibull shape parameter, dimensionless
C	chuck to bushing distance, inches
C1	first polynomial coefficient, dimensionless
C2	second polynomial coefficient, dimensionless
C3	third polynomial coefficient, dimensionless
C4	fourth polynomial coefficient, dimensionless
D	difference in the lengths of the two broken pieces of the specimen, inches
d	specimen and shaft diameter, inches
E	modulus of elasticity, psi
F	cumulative probability of failure, dimensionless
f	probability density function, dimensionless
h	height of specimen loop, inches
i	rank of the ordered observation, dimensionless
j	rank of the stress level, dimensionless
L	wire external to chuck and bushing, inches
Lt	total length of the specimen, inches
L{DATA}	likelihood operator
N	number of cycles to failure, dimensionless
n	number of specimens in sample, dimensionless
P	probability of failure, dimensionless
R	reliability, dimensionless
$\bar{R}(x_i)$	reliability associated with ith failure, dimensionless
Rmin	minimum radius of curvature of fatigue specimen, inches
S	stress, psi
s	load induced stress, psi
W_d	uncertainty in diameter, inches
W_E	uncertainty in modulus of elasticity, psi
W_L	uncertainty in length, inches
W_s	uncertainty in stress, psi
x	variate
x_o	Weibull guaranteed life parameter
Γ	gamma function operator
θ	Weibull characteristic life parameter
μ	mean
σ	standard deviation

Introduction

Fatigue data traditionally have been presented graphically as an S-N diagram. The general technique used in determining the S-N curve is to test a few specimens at each stress level, find the average number of cycles to failure for each stress level, and then plot the log of cycles to failure versus the log of the stress at failure as shown in Figure 1.

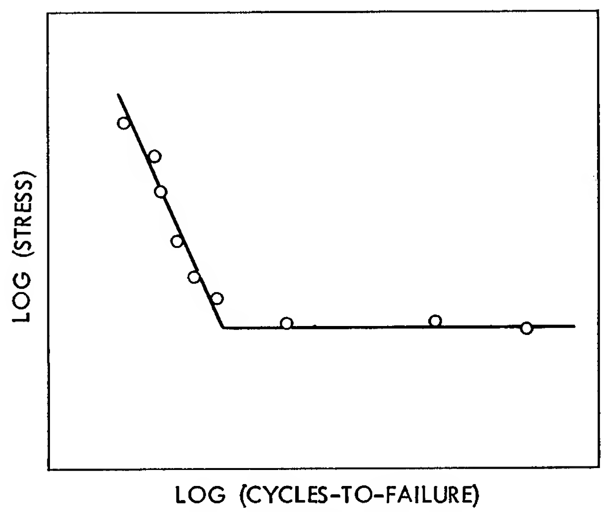


Fig. 1. Conventional S-N curve.

A much better way of representing the data is with a three-dimensional S-N-density curve showing both cycles-to-failure distributions and stress-to-failure distributions as shown in Figure 2.

Most experimenters in the past have chosen either the normal or the log-normal distributions to approximate the cycles-to-failure distributions depending upon how far the data points deviate from a straight line when plotted on probability paper. When data points are plotted on normal probability paper, the agreement with the theoretical normal distribution function is revealed by the extent to which the points fall along a straight line. If the agreement is poor and the distribution appears to be positively skewed, then the log-normal distribution would probably yield better fidelity.

The American Society for Testing and Materials, committee E-9,[1] states that some fatigue tests, particularly those made in the finite life range of an S-N curve, may yield approximately normal distributions of

cycle life, but generally require a transformation to log-cycle life. However, others do not yield normal distributions, even after various transformations are performed on the data.

Juvinall[2] states that, while experimental points fall reasonably close to straight lines on log-normal paper, justification is not indicated for making precise predictions of reliability according to log-normal relationships.

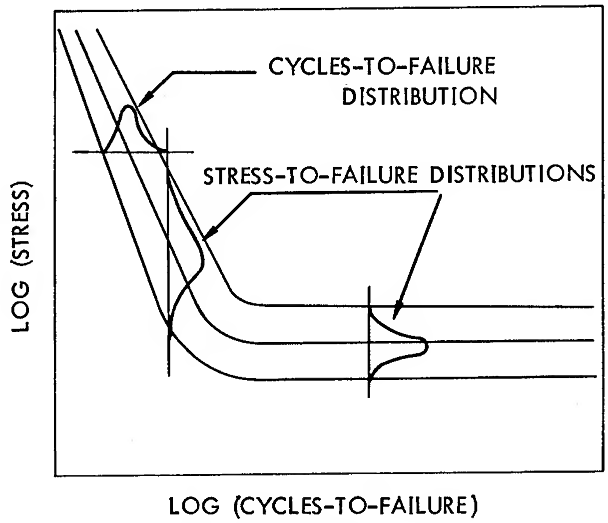


Fig. 2. S-N-density curve.

Kececioglu, Smith, and Felsted[3] concluded from the results of steel wire fatigue tests that cycles-to-failure data at each stress level, for stresses significantly above the endurance strength, were best represented by a log-normal distribution. However, in tests the following year on SAE 4340 wire specimens of 0.0625 inch diameter, Rececioglu and Haugen[4] found that at high stress levels, the normal and the log-normal distributions represent the data almost equally well, whereas at the lower stress levels, the log-normal distribution provided a superior fit.

Very little work has been done to date on stress-to-failure distributions. Stulen, Cummings, and Schulte,[5] analyzing fatigue data obtained by various investigations, have concluded that stress-to-failure distributions have a reasonably normal (rather than log-normal) distribution.

Kececioglu, Smith, and Felsted,[3] in analyzing data from both cold drawn steel wire and 7075-T6 aluminum wire, concluded that as the life of the specimen increases, the mean strength decreases, while the standard deviation appears to increase slightly for the steel specimen and significantly for the aluminum specimen. Furthermore, the coefficient of skewness is negative for most of the stress-to-failure distributions, indicating a normal distribution fit is superior to the log-normal.

Data Acquisition

The material used in this investigation was straight music wire 0.0242 inches in diameter. The size variation and physical properties as given by the manufacturer are:

Wire diameter - - - - - - 0.0240 inches
Allowable variation - - - $\pm$0.0003 inches
Tensile strength - - - - - 341,000 - 371,000 psi
Yield strength - - - - - - 225,000 psi (estimated)
Modulus of elasticity - - 30 x 10^6 psi.

Chemical composition:

Carbon - - - - - - - - - - 0.70 to 0.90 %
Manganese - - - - - - - - - 0.20 to 0.40 %
Phosphorus - - - - - - - - 0.025 % max.
Sulphur - - - - - - - - - 0.025 % max.
Silicon - - - - - - - - - 0.12 to 0.25 %

Wire Fatigue Tester

The fatigue tester used to conduct this investigation was a rotary beam fatigue tester, model 802, manufactured by the Hunter Spring Company. The machine is of the large-deflection, slender-column variety. The specimen is looped a complete 180 degrees between the chuck and the bushing and rotated to give completely reversed bending.

The machine consists of a motor driven chuck and a magnetic bushing. The bushing can be positioned in any one of nine holes in the bushing support spaced at one-inch intervals. Fine adjustment is accomplished by horizontal movement of the bushing support which is attached to a micrometer dial indicator calibrated in thousandths of an inch.

A 1/50 horsepower synchronous motor, operating at 3600 rpm, turns the chuck while an electric time meter graduated in tenths of a minute is used to register elapsed time. An electronic cut-off circuit, which will operate under a contact resistance from 100,000 to over 20,000,000 ohms, is used to turn off the motor and timer when the specimen fails. A wire form, connected to the cut-off circuit, is mounted on a movable magnetic base and positioned so the wire specimen, upon breaking, makes contact with the wire form and activates the cut-off circuit.

Two wire guides, used to minimize excessive vibration and sag in the specimen, are mounted on movable magnetic bases. The guides, which support the wire specimen, are positioned outside of the region of maximum stress.

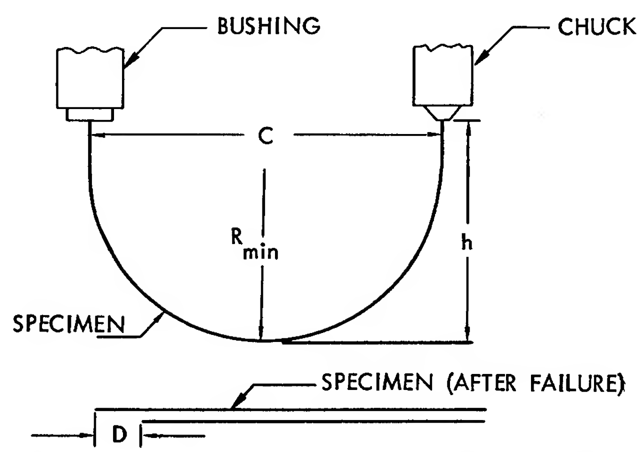


Fig. 3. Physical significance of testing machine parameters.

The following equations express the relationships between the different parameters:

$C=1.198\ Ed/S$	inches	(1)
$L=2.19\ C$	inches	(2)
$h=0.835\ C$	inches	(3)
$R_{min}=0.417\ C$	inches	(4)
$L_t=L+0.75$	inches	(5)

The derivations of these equations can be found in a paper by F. A. Volta.[6]

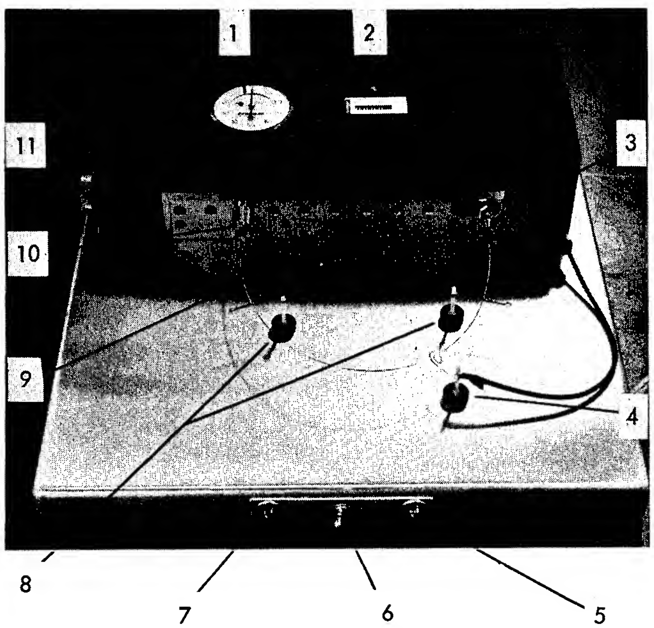
Set-up Procedure and Operation of the Tester

The value of stress is selected and the chuck-to-bushing distance C and specimen length L are calculated according to Equations (1) and (2), respectively. An

additional allowance of 0.75 inches is made for the portions of the specimen gripped in the chuck and the bushing. The total length of the specimen is then denoted as Lt in Equation (5). The chuck-to-bushing distance is set by inserting the bushing in the appropriate hole and turning the adjusting knob until the proper decimal fraction is indicated on the dial indicator.

The wire is carefully wiped with a clean cloth to remove any deposits present on the surface and then fastened securely in the chuck. The other end is held in the bushing due to the force of a permanent magnet located directly behind the bushing. The wire guides and electronic cut-off contact are positioned outside the region of maximum stress as shown in Figure 4.

The time meter is set to zero and the power switch turned to the on position to energize the motor, timer, and both pilot lights. Upon failure of the specimen, the motor, time meter, and the operate pilot light are de-energized. The time meter reading is recorded and the specimen removed from the tester. The electronic cut-off mechanism is reset by turning off the power switch.



1. DIAL INDICATOR
2. TIME METER
3. CHUCK
4. MAGNETIC BASE CUT-OFF POST
5. OPERATE PILOT LIGHT
6. ON-OFF SWITCH
7. POWER PILOT LIGHT
8. MAGNETIC BASE GUIDES
9. SPECIMEN
10. MOVABLE MAGNETIC BUSHING
11. ADJUSTMENT KNOB

Fig. 4. The Hunter rotating-beam fatigue machine.

Testing Procedure

The upper bound for the operating stress was chosen as the yield stress of the specimen. The endurance limit, which was determined experimentally, was taken as the lower bound.

For each stress level S, the parameters C, L, Lt, h, and Rmin were calculated and the results tabulated in Table 1. Twenty specimens were tested at each of the 26 stress levels for a total of 520 specimens.

The specimens were chosen at random to minimize temperature variation effect and were tested as described in the section on set-up procedure and operation of the tester. Upon failure, 0.02 minutes were subtracted from the elapsed time indicated on the time meter due to over-run and the time was recorded. The number of cycles-to-failure was calculated by multiplying the time by 3600 rpm. The difference D in the lengths of the two broken pieces of the specimen was measured and the D/L ratio recorded along with the percent of maximum stress as given in Figure 5. The number of cycles-to-failure were then ordered in ascending order. The rank, number of cycles-to-failure, D/L, percent of maximum stress, and the ambient temperature for each specimen are given in Table 2.

Table 1. Testing machine set-up parameters.

Stress, psi	C, inches	L, inches	Lt, inches	h, inches	Rmin, inches
100000	8.697	19.047	19.797	7.262	3.627
105000	8.283	18.140	18.890	6.917	3.454
110000	7.907	17.316	18.066	6.602	3.297
115000	7.563	16.563	17.313	6.315	3.154
120000	7.248	15.873	16.623	6.052	3.022
125000	6.958	15.238	15.988	5.810	2.901
130000	6.690	14.652	15.402	5.585	2.790
135000	6.443	14.109	14.859	5.380	2.687
140000	6.212	13.605	14.355	5.187	2.591
145000	5.998	13.136	13.885	5.009	2.501
150000	5.798	12.698	13.448	4.842	2.418
155000	5.611	12.289	13.039	4.685	2.340
160000	5.436	11.905	12.655	4.539	2.267
165000	5.271	11.544	12.294	4.401	2.198
170000	5.116	11.204	11.954	4.272	2.133
175000	4.970	10.884	11.634	4.150	2.072
180000	4.832	10.582	11.332	4.035	2.015
185000	4.701	10.296	11.046	3.926	1.960
190000	4.578	10.025	10.775	3.822	1.909
195000	4.460	9.768	10.518	3.724	1.860
200000	4.349	9.524	10.274	3.631	1.813
205000	4.243	9.291	10.041	3.543	1.769
210000	4.142	9.070	9.820	3.458	1.727
215000	4.045	8.859	9.609	3.378	1.687
220000	3.953	8.658	9.408	3.301	1.649
225000	3.866	8.466	9.215	3.228	1.612

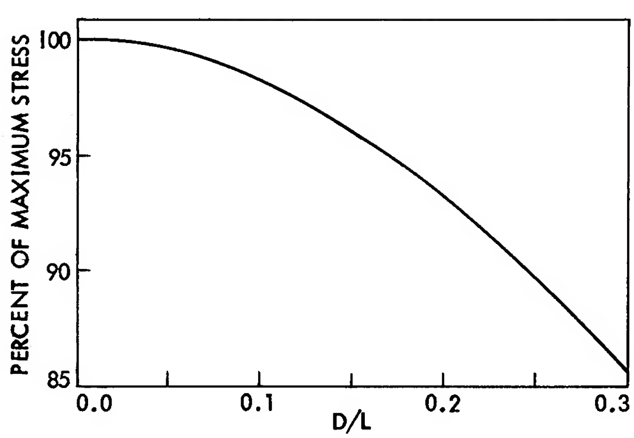


Fig. 5. Percent of maximum stress versus D/L ratio.

Table 2. Sample Fatigue data.

Rank	Cycles to failure	D/L	% of maximum stress	Temp., °F
	Maximum stress = 100000 psi			
1	53028	0.0557	99.469	78
2	61560	0.0368	99.768	78
3	64440	0.0525	99.527	82
4	81720	0.0336	99.805	80
5	83628	0.0347	99.794	78
6	89280	0.0252	99.891	80
7	90792	0.0557	99.469	80
8	130068	0.0651	99.274	80
9	146628	0.0105	99.981	83
10	164268	0.0872	98.703	79
11	174708	0.0646	99.286	82
12	990648	0.0357	99.781	78
*13	1367927	0.0000	100.000	77
*14	1750247	0.0000	100.000	80
*15	1763927	0.0000	100.000	79
*16	2148047	0.0000	100.000	78
*17	2296727	0.0000	100.000	77
*18	2861927	0.0000	100.000	77
*19	4319927	0.0000	100.000	80
*20	15757110	0.0000	100.000	77
	Maximum stress = 105000 psi			
1	26280	0.0386	99.744	76
2	53496	0.0617	99.347	73
3	57960	0.1433	96.531	78
4	64188	0.0733	99.080	80
5	68220	0.0882	98.672	82
6	68436	0.0171	99.950	79
7	69768	0.1577	95.817	76
8	71064	0.0320	99.824	79
9	74772	0.1329	97.012	82
10	76068	0.0932	98.520	78
11	77976	0.0198	99.932	78
12	79380	0.0204	99.929	79
13	88956	0.0474	99.614	78
14	89856	0.0943	98.485	78
15	90108	0.1764	94.791	81
16	93600	0.0882	98.672	80
17	108864	0.1158	97.723	76
18	110988	0.0138	99.967	77
19	114732	0.1080	98.014	80
20	118116	0.1632	95.526	78

*Indicates that specimen did not fail.

Presentation of Results

Smoothing of the cycles-to-failure data given in Table 2 was necessary so that stress-to-failure data could be generated. The approach is similar to the one shown in the ASTM Special Technical Publication No. 121.[7]

The relative cumulative frequency for each failure at the jth stress level S(j) is given by Equation (6):

$$P(i,j) = i/(n + 1) \tag{6}$$

where $P(i,j)$ = Relative cumulative frequency of the ith failure at the jth stress level,
n = Total number of specimens in the sample space,
i = Rank of the ordered observation, and
j = Rank of the stress level.

P(i,j) is also known as the probability of failure at life measure N(i,j) where N(i,j) is the number of cycles to failure for the ith sample at the jth stress level.

If N(1,j) is regressed on S(j) for all values of j, then a least squares curve is obtained for all the specimens of rank one on the S-N diagram. If this is done for the remaining ranks, then a total of n constant probability curves will be formed. These curves are then plotted on an S-N diagram to give a P-S-N diagram as in Figure 6.

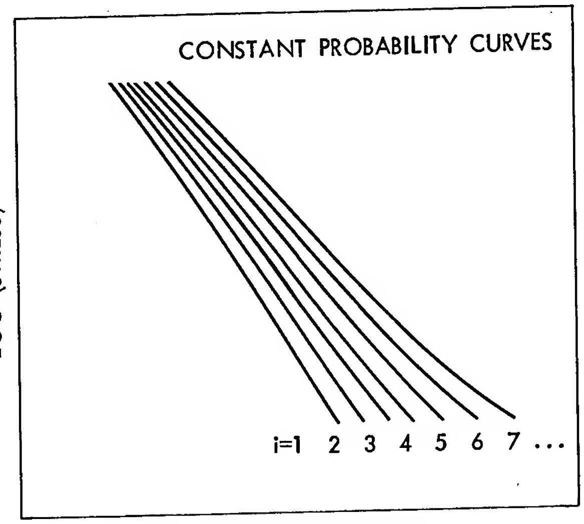


Fig. 6. P-S-N diagram.

The data at the 100,000-psi stress level were not included in the data sets used for the curve fits because there were a number of specimens that did not fail at this stress level. It was found that the best fit to the data was obtained using third-degree polynomials, the coefficients of which are given in Table 3. The form of the polynomial is given in Equation (7):

$$y = C1 + C2 \cdot x + C3 \cdot x^2 + C4 \cdot x^3 \tag{7}$$

where $x = \ln(S)$ and $y = \ln(N)$.

Table 3. Polynomial coefficients.

Rank	C1	C2	C3	C4
1	-3589.553	893.0386	-73.56918	2.011050
2	4846.617	-1212.038	101.5240	-2.842877
3	5838.238	- 954.7173	79.55891	-2.224159
4	5583.355	-1392.193	115.2088	-3.241850
5	6420.984	-1601.722	133.6805	-3.727474
6	5559.160	-1386.413	115.7550	-3.230132
7	5417.602	-1349.192	112.4978	-3.135243
8	5460.613	-1359.738	113.3634	-3.159021
9	5897.488	-1467.947	122.2973	-3.404881
10	5377.094	-1336.214	111.1829	-3.092319
11	5502.113	-1368.033	113.8837	-3.168757
12	5011.988	-1244.347	103.4851	-2.877494
13	5926.211	-1469.988	122.0418	-3.385992
14	6484.895	-1611.260	133.9498	-3.720575
15	6723.172	-1671.041	138.9485	-3.859871
16	6683.246	-1659.547	137.8668	-3.826431
17	8476.063	-2106.891	175.0687	-4.857531
18	8022.160	-1990.444	165.1177	-4.574261
19	8986.156	-2231.535	185.2123	-5.132412
20	9904.270	-2456.437	203.5676	-5.631505

Both smoothed cycles-to-failure and stress-to-failure data were generated using the polynomial coefficients in Table 3 at the following stress levels and cycles-to-failure levels:

Stress levels: 105000, 115000, 125000, 135000, 145000, 155000, 165000, 175000, 185000, 195000, 205000, 215000, and 225000 psi.

The following cycles-to-failure levels were chosen on a logarithmic basis so that the entire P-S-N diagram would be covered uniformly. The highest level chosen was 26903 since extrapolation would occur beyond this point.

Cycles-to-failure levels: 2981, 3641, 4447, 5432, 6634, 8103, 9897, 12088, 14764, 18033, 22026, 26903.

Weibull distributions were fitted to the generated cycles-to-failure and stress-to-failure data by both a least squares reliability curve fitting technique and the maximum likelihood technique.

Least Squares Reliability Curve Fitting Technique

The mean and the standard deviation were calculated for each set of data generated from the polynomial curves. Knowing these two values and choosing a value for the Weibull guaranteed life x_o, Equations (25) and (26) in Appendix A[8] can be solved for the values of b and θ.

The reliability associated with the ith failure event at life measure x_i is given by Equation (8):

$$\bar{R}(x_i) = (n + 1 - i)/(n + 1). \quad (8)$$

This expression presents a very small amount of bias in the case of machine failures. The reliability for a Weibull distribution is denoted as R(x) as given by Equation (27) in Appendix A[8]. The problem at hand is to find the value of x_o that will minimize the following expression:

$$\text{Sum} = \sum_{i=1}^{n} \left[\bar{R}(x_i) - R(x_i)\right]^2 . \quad (9)$$

This task is easily accomplished by use of a one-dimensional search on a digital computer. The task could also be accomplished by plotting the reliabilities on Weibull probability paper and adjusting the value of x_o such that the deviation of the data points from a straight line will be minimized. The resulting Weibull parameters as calculated by the least squares reliability curve fitting method are summarized in Tables 4 and 5.

Table 4. Weibull parameters for cycles-to-failure distributions (least squares method).

Stress level, psi	μ	σ	b	θ	x_o
105000	76648	21876	2.167	50793	31692
115000	47031	9953	2.996	30632	19696
125000	31861	5290	3.774	19801	13983
135000	23044	3215	4.408	13729	10522
145000	17411	2189	4.780	10015	8233
155000	13545	1627	4.879	7577	6593
165000	10740	1284	4.758	5854	5381
175000	8619	1050	4.495	4560	4455
185000	6964	872	4.198	3572	3719
195000	5644	725	3.910	2798	3109
205000	4575	598	3.718	2211	2577
215000	3702	487	3.589	1747	2129
225000	2987	390	3.528	1378	1746

Table 5. Weibull parameters for stress-to-failure distributions (least squares method).

Cycles-to-failure level	μ	σ	b	θ	x_o
2981	224577	6205	5.754	33292	193763
3641	215322	6355	5.721	33928	184004
4447	205951	6381	5.582	33352	175206
5432	196497	6295	5.476	32361	166629
6634	187034	6110	5.508	31570	157951
8103	177641	5858	5.690	31122	148915
9897	168440	5579	6.101	31485	139263
12088	159565	5326	6.700	32611	129130
14764	151140	5150	7.584	35161	118166
18033	143255	5100	9.114	40994	104413
22026	135956	5226	12.769	57033	81131
26903	129232	5620	42.296	190978	- 59198

The cycles-to-failure and stress-to-failure distributions calculated according to the parameters specified in Tables 4 and 5 are shown in Figures 7, 8, and 9. An S-N curve representing the locations of the means of the smoothed cycles-to-failure distributions is plotted along with the means of the original cycles-to-failure distributions in Figure 10.

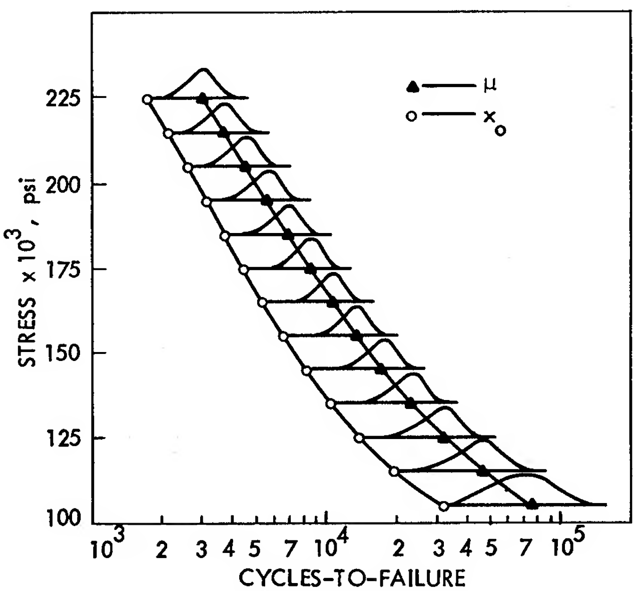


Fig. 7. Semilogarithmic S-N-density diagram of cycles-to-failure distributions.

Maximum Likelihood Technique

For most parametric estimation problems, the method of estimation called the method of maximum likelihood is the most efficient method available. The method used for the estimation of Weibull parameters is described below.

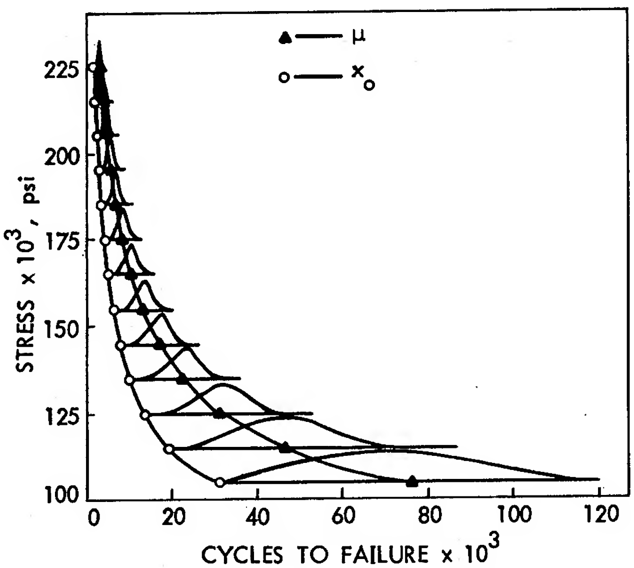


Fig. 8. Cartesian coordinate S-N-density diagram of cycles-to-failure distributions.

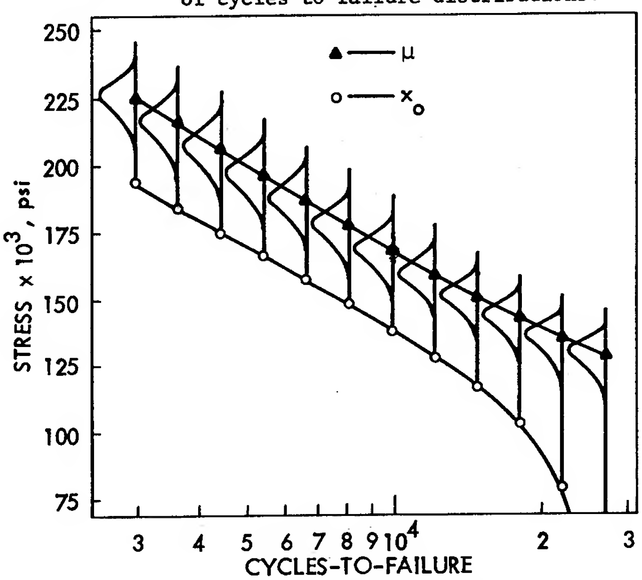


Fig. 9. Semilogarithmic S-N-density diagram of stress-to-failure distributions.

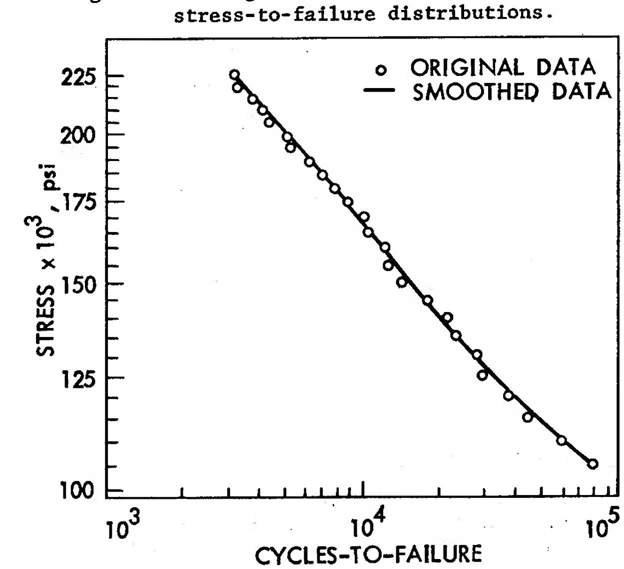


Fig. 10. Comparison of the means obtained from the original data with the means obtained from the smoothed data.

Given a set of data

$$\{DATA\} = \{x_1, x_2, x_3, \ldots, x_n\} \tag{10}$$

the probability of observing x_i is

$$P(x_i) = f(x_i)\Delta x \tag{11}$$

where $f(x)$ is the probability density function.

The probability of observing the data set is

$$P\{DATA\} = [f(x_1)\Delta x]\,[f(x_2)\Delta x]\,\ldots\,[f(x_n)\Delta x] \tag{12}$$

$$P\{DATA\} = \overline{\Delta x^n} \prod_{i=1}^{i=n} f(x_i). \tag{13}$$

The likelihood function is defined as

$$L\{DATA\} = \ln\left[\frac{P\{DATA\}}{\Delta x^n}\right] = \ln \prod_{i=1}^{i=n} f(x_i)$$

$$= \sum_{i=1}^{i=n} \ln f(x_i) \tag{14}$$

To find the values of b, θ, and x_c which make L a maximum, a multidimensional search is employed which will vary the above parameters in a systematic manner until the maximum value of L is reached. At this point the search will terminate and return the best estimates for b, θ, and x_o.

The results of the computer output for the best estimates of the Weibull parameters are given in Tables 6 and 7. The Weibull shape parameter b as a function of stress level is shown in Figure 11. Figure 12 shows b as a function of cycles-to-failure level.

Table 6. Weibull parameters for cycles-to-failure distributions (maximum likelihood method).

Stress level, psi	b	θ	x_o
105000	2.980	67532	16332
115000	3.618	35773	14794
125000	3.842	19597	14066
135000	4.545	13751	10512
145000	4.936	10020	8238
155000	5.026	7580	6597
165000	4.885	5856	5380
175000	4.598	4562	4456
185000	4.271	3574	3715
195000	3.968	2799	3110
205000	3.438	2044	2739
215000	3.006	1483	2379
225000	2.450	993	2102

Table 7. Weibull parameters for stress-to-failure distributions (maximum likelihood method).

Cycles-to-failure level	b	θ	x_o
2981	5.812	33315	193762
3641	5.785	33905	183963
4447	5.660	33330	175167
5432	5.568	32360	166628
6634	5.605	31550	157915
8103	5.812	31103	148881
9897	6.252	21467	189231
12088	6.900	32610	129129
14764	7.883	35144	118137
18033	9.555	40978	104412

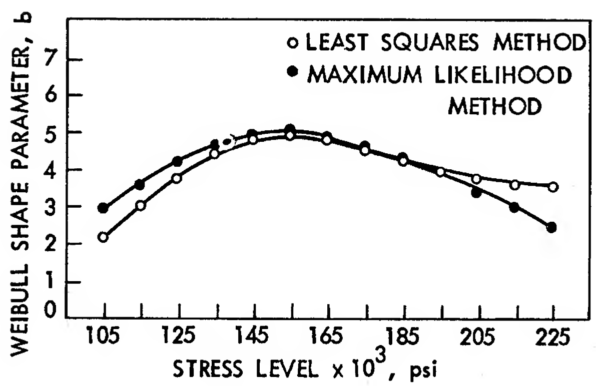


Fig. 11. Weibull shape parameter versus stress level.

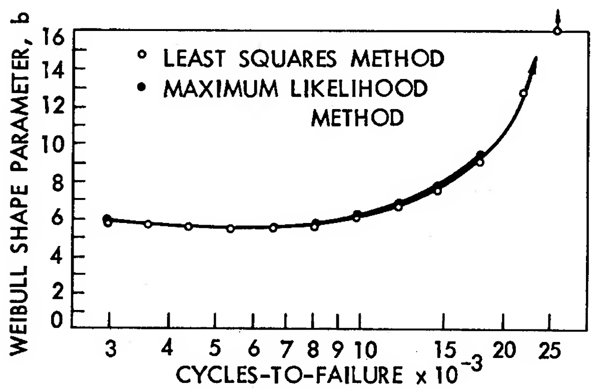


Fig. 12. Weibull shape parameter versus cycles-to-failure level for stress-to-failure destributions.

Conclusions

As a result of this study, the six following conclusions appear to be justified.

1. The shape of the cycles-to-failure distribution is not constant over the entire stress range.

2. Based on the Weibull shape parameters from both the least squares method and the maximum likelihood method, a normal distribution would represent cycles-to-failure data better at medium stress levels than a log-normal distribution. However, the log-normal distribution appears to be better at both the high and low stress levels than the normal distribution.

3. The shape of the stress-to-failure distributions is not constant over the entire range of cycles-to-failure.

4. Based on the Weibull shape parameters from both the least squares method and the maximum likelihood method, stress-to-failure distributions are definitely not log-normal and are not even represented well by the normal distribution.

5. The Weibull distribution possesses the necessary flexibility required to better represent fatigue distributions. An experimenter could make a serious error in overlooking the possibility that neither the normal nor the log-normal distributions fit the data properly.

6. The results of this investigation should not be generalized for fatigue distributions of other types of materials without further investigation.

Acknowledgment

This work was supported by the Engineering Research Institute, Iowa State University, Ames, Iowa 50010.

Bibliography

1. American Society for Testing and Materials. A guide for fatigue testing and the statistical analysis of fatigue data. ASTM Special Technical Publication No. 91-A. 1963.

2. Juvinall, Robert C. Stress, strain, and strength. McGraw-Hill Book Co., New York. 1967.

3. Kececioglu, Dimitri B., Smith, Richard E., and Felsted, Ernest A. Distributions of strength in simple fatigue and the associated reliabilities. Annals of Assurance Sciences of the 8th Reliability and Maintainability Conference, Denver, Colorado. July, 1969.

4. Kececioglu, Dimitri, and Haugen, Edward B. Fatigue reliability design data for dynamic and rotary machinery. American Society of Mechanical Engineers, Space Technology and Heat Transfer Conference, ASME Paper No. 70-Av/SpT-36, Los Angeles, California. June, 1970.

5. Stulen, F. B., Cummings, H. N., and Schulte, W. C. Preventing fatigue failures, Part 5, Calculating fatigue strength. Machine Design, Vol. 33. June 22, 1961.

6. Volta, F. A. New wire fatigue testing. Iron Age, Vol. 162, No. 7:78-81. August, 1948.

7. American Society for Testing and Materials. Symposium on statistical aspects of fatigue. ASTM Special Technical Publication No. 121. 1951.

8. Wagner, G. R., Cycles-to-failure and stress-to-failure Weibull distributions in steel wire fatigue Master of Science Thesis, Iowa State University, Ames, Iowa. 1972.

Supplementary References

Cazaud, R. Fatigue of metals. Chapman and Hall, Ltd., London. 1953.

Davis, H. F., Troxell, G. E., and Wiskocil, C. T. The testing and inspection of engineering materials. McGraw-Hill Book Co., New York. 1964.

Forrest, P. G. Fatigue of metals. Pergamon Press, New York. 1962.

Forsyth, P. J. E. The physical basis of metal fatigue. Blackie and Son, Ltd., London and Glasgow. 1969.

Goodman, Irwin F. Bibliography on the Weibull distribution. Symposium for reliability and maintainability technology for mechanical systems. American Ordnance Association, Washington, D. C. April 27, 1972.

Gough, H. J. The fatigue of metals. Ernest Benn, Ltd., London. 1926.

Grover, H. J., Gordon, S. A., and Jackson, L. R. Fatigue of metals and structures. Department of the Navy, U.S. Government Printing Office, Washington, D.C. 1954.

Kennedy, A. J. Processes of creep and fatigue in metals. John Wiley and Sons, Inc., New York. 1963.

Sines, G., and Waisman, J. L. Metal fatigue. McGraw-Hill Book Co., Inc., New York. 1959.

Timoshenko, Stephen P. History of strength of materials. McGraw-Hill Book Co., New York. 1953.

Weibull, W. Fatigue testing and analysis of results. Pergamon Press, New York. 1961.

SYSTEM ENGINEERING ASPECTS OF THE MAN - MACHINE INTERFACE

By: Edward T. Parascos INDEX SERIAL NUMBER – 1105

Quality Assurance and Reliability Consultant
Consolidated Edison Company of New York, Incorporated
4 Irving Place
New York, N.Y. 10003

Summary

Man-machine interfaces are increasing in complexity in the modern technology of today. In the past, Human Factor Engineering considerations played a relatively minor role in the design of equipment. Human Factor design checklists were used to assess the ability to operate, a newly conceived system and usually after its design was complete. In this age of complexity this approach is no longer adequate, it never really was.

As machines and interaction between machines became more complex, a new engineering discipline emerged, Systems Engineering. This new discipline embraced at least the following functions;

Operations Research
Statistics
Electronic Data Processing
Cost Accounting and
Design Engineering

in order to systemize, optimize and symplify the design of complex equipments.

Similarly, as the man-machine interface also became complex, a new form of Human Factor Engineering emerged, one that was System Engineering oriented.

A comparison is made here-in, between the early Human Factor approach and the current Human Factor approach with its emphasis on Systems Engineering. MIL-H-46855, Human Engineering Requirements for Military Systems, Equipments and Facilities, is reviewed. This specification looks like, acts like and in reality is a Systems Engineering Specification. The following elements of MIL-H-46855 are illustrated:

Function and Time Line Analysis
Operation Sequence Diagrams
Crew Loading Analysis
Symbols, etc.

Finally the results of applying these current Human Factor requirements on "State of the art" LASER scanning and recording subsystems, are discussed.

Human Factor Engineering is a function based on the concept that man is a vital component in many systems. There are many elements that are a part of Human Factor Engineering. This paper will not try to cover every aspect of Human Factor Engineering, but the references listed at the end of this paper will be of use towards this end.

Of all the elements concerned with the Human Factor discipline, there are four (4) which provide the basic background. These are:

Time and motion study
Anthropology
Psychology
Systems Engineering

To digress, momentarily, look at Figure number 1. How many of you see a beauty? How many of you see something else? How many of you see both? How many of you see nothing? The point of this example is that each and every one of you sees what he wants to see. This inconsistency in response to a fixed stationary black and white object is very baffling to the design engineer who has been trained in providing a single or limited set of solutions to a set of parameters.

Human Factors Engineering brings to the man-machine interface the idea that there are many solutions to a problem, as many perhaps as there are individuals who will operate this machine. The Human Factor specialists in this situation assist in choosing from the many solutions, the optimum one.

In the past Human Factor Engineering specifications emphasized only the Anthropological aspects of design. This class of specifications terminated with MIL-STD-803 and MIL-STD-1472. These specifications in one way or another contained the following type of statement, which is from MIL-STD-1472.

"The equipment shall represent the simplest design consistent with functional requirements and expected service conditions. To the maximum extent possible it shall be capable of operation, maintenance and repairs by personnel with a minimum of training." Such a statement has a different meaning to each engineer who reads it. It is what is known as a "Motherhood" statement. These

specifications were useful but only to a limited extent. They encouraged design of displays which would effectively present information to the operator and the design of controls which would permit him to manipulate the system with equal effectiveness. This period in Human Factor Engineering history was known as the "Knob and dial" era.

In February of 1968 a new type of Human Factor Specification was released, MIL-H-46855. The scope of this specification which follows sets the standard for a new era:

"This specification establishes and defines the general requirements for applying the principles and criteria of human engineering to the concept formulation, definition, and acquisition of military systems, equipment and facilities. The requirements include the work to be accomplished or subcontracted by the contractor in effecting an integrated human engineering effort. Compliance with these requirements forms the basis for including human engineering during proposal preparation and data reporting by the contractor (e.g. such items as flow charts, functional allocation tables, operational sequence diagrams, link analysis, task descriptions, etc.) as specified by the Contract."

Though this specification encompasses the "Knob and dial" concepts, since it lists as applicable documents MIL-STD-1472, it introduces the disciplines of Time and Motion Study," Systems Engineering and Psychology.

This specification was imposed on the Joint Services In-Flight Data Transmission System (JIFDATS). JIFDATS provided a near real time transfer of reconnaissance data through a microwave data transmission system, from a sensor equipped aircraft to a surface terminal, either by a direct link or through an airborne relay when the sensor aircraft is beyond the radio line of sight of the surface terminal. Surface terminals fell in two broad categories, ship and land based. Equipment were designed for common usage with either surface terminal concept. The land-based terminal was also to be self supporting, mobile, and air-transportable for maximum tactical flexibility.

CBS Laboratories was a major subcontractor providing an In Flight Photographic Processor and Scanning System (IPPS) in the Sensor Aircraft and Photographic Recorder Processor Viewer (PRPV) in the surface terminal. Figure 2 presents CBS Laboratories contribution to the JIFDATS Program. The IPPS equipment would operate on film fed to it from a photographic sensor (camera). Its unique dry processor would develop the latent image on the exposed film and feed the processed film to the Laser Scanner which was an integral portion of IPPS. The Scanner would scan the film and convert the photographic image into an electronic one. The electronic image would then be transmitted directly to the surface terminal or indirectly to the surface terminal after being relayed by the relay aircraft. The PRPV in the surface terminal would through its LASER Recording section record the acquired electronic image photographically on raw film and feed the film to the Photographic Processing section which would develop the exposed film, dry it and pass it through a viewer for interpretation by a photo interpreter. This entire operation took under 10 minutes from the time the IPPS received its first frame until the photo interpreter viewed this same frame.

A Mission Scenario and Functional Flow diagrams for JIFDATS operation were supplied to each major subcontractor. A Mission Scenario is a thoroughly complete narrative of each and every facet of a tactical military mission. Functional flow diagrams are block diagrams of every operation required in a tactical mission. Figure 3 is a small portion of one of the ten (10) Functional Flow diagrams supplied to CBS Laboratories immediately after Contract award.

Utilizing the Mission Scenario and the Functional flow diagram supplied by the prime contractor, Human Factor Engineering analyses were conducted on the IPPS and the PRPV based on the following requirements of MIL-H-46855. Figure 4 presents the Time and Motion Study symbology used in these analyses.

Function Analysis (Paragraph 30.3.2 of MIL-H-46855)

"A detailed time line analysis of the system response requirements as related to system mission is prepared for normal and degraded system operation. The analysis includes all modes of operation, secondary operation and the probability of using each mode of operation or system state. The amount of degration which can occur without affecting mission success is determined for each mode of operation and for the system as a whole. This analysis begins with a block diagram of functions required to complete the mission and the modes involved. All functions are contained in two categories, decision functions and action functions. Decision functions are reduced until no further binary (Yes, No) decision is possible. Reiterative decisions or frequently utilized segments may be indicated as subroutines."

Applying this requirement to the IPPS and the PRPV resulted in thirty seven (37) separate analyses one of which is illustrated in Figure 5.

Allocation of Functions (Paragraphs 3.2.1.1.3 & 30.3.3)

"Tables showing allocations of functions and presenting rationales are prepared for those functions requiring critical human involvement, and those which should be machine-implemented."

Applying this requirement resulted in thirty four (34) tables for the IPPS and the PRPV one of which is illustrated in Figure 6.

Operation Sequence Diagrams (OSD's) (Paragraph 30.3.4)

"An Operation Sequence Diagram is a technique of plotting relative to time (actual or sequential), the flow of information, data or energy through an operationally defined system using standard symbols relating to actions taken (inspections, data transmittal, data receipt, data storage, or decisions), as that data, information or energy is manipulated internally in the system by defined men and equipment. Once functions are allocated to either man or machine and major subsystems are identified, OSD's are prepared on a time base for typical system operation."

Applying this requirement for the IPPS and the PRPV resulted in fifty five (55) OSD's, one of which is illustrated in Figure 7.

Operator/Maintainer Information Requirements (Paragraph 30.3.5)

"Using OSD's and other relevant information, tables of information requirements are prepared for all operator/maintainer positions. These tables indicate all the inputs, processing, and outputs for these positions, including quantitative expression of load, accuracy, rate, and time delay. Comprehensive information on these factors is developed to provide an adequate basis for defining control, display, and communication requirements."

Thirty two (32) tables listing operator information requirements were required for IPPS and PRPV. One of these tables is illustrated in Figure 8.

Control, Display, and Communication Requirements (Paragraph 30.3.6)

"OSD's and information requirements data is used to generate control, display, and communication requirements for each operator/maintainer position. The definition of requirements must be comprehensive and complete so as to allow direct translation into hardware configurations and software programs appropriate to the man-machine interface."

Eleven (11) Control, Display, and Communication Requirement tables were developed for the IPPS and the PRPV to satisfy this requirement. One of these tables is illustrated in figure 9.

Operator/Maintainer Task Descriptions (Par. 30.3.7)

"Task-related data is extracted from the OSD's and requirements summaries. This data is compiled in preliminary operator/maintainer procedurally-oriented task descriptions for later use in developing procedures documents, personnel planning and system testing. Wherever there is critical human involvement it is noted, together with the consequences of error or time delay. The analyses include operator interaction where more than one operator is included. All missions and phases are included. Operating modes are analyzed and provision for analytical treatment of the less than 100% reliabilities of operator and hardware are also made."

The IPPS and PRPV sub-systems provided thirty-three (33) Operator/Maintainer Task Description to the total JIFDATS System. One of these is illustrated in figure 10.

Crew Loading Analysis (Par. 30.3.8)

"A time profile analysis of operator work load is prepared. Supporting evidence for action times are presented. Where possible, distribution function times are considered in the analysis. A condition of maximum operator work load based on operator action times are determined and prepared for the primary mission and any phase where the operator loading exceeds 75% utilization. The influence of training and retention is also analyzed."

Twenty four (24) such analyses were developed for the IPPS and PRPV one of which is illustrated in figure 11.

Personnel Planning Information (Par. 30.3.9)

"Using all the descriptions, summaries and analyses just presented human factor personnel prepare a first cut summary of personnel planning information, indicating the level of operator/maintainer ability

required and profile of skills and knowledges needed for each operator/maintainer in the system.

Special skills, knowledges and selection requirements related to critical human involvement are noted and documented."

CBS Laboratores was not required to perform this task, since the prime contractor was coordinating the entire requirements. Conducting these types of human Factor Engineering analyses meant working closely with the design personnel. This had a dramatic effect on both types of individuals. The Human Factor analyst's for the first time felt that they were contributors at the conceptual phase. They are usually brought in after the system is completely designed and are ignored if they suggest major or sometimes even minor changes.

The design engineers changed even more dramatically. MIL-H-46855 was imposed during the proposal phase. Human Factor Engineering Plans were to be completed within a month after contract award. Preliminary Analyses had to be submitted within 45 days after contract award. Human Factor Engineering personnel had to participate in all Design Reviews and had sign off authority on all drawings having an impact on the man-machine interface.

Design personnel discovered for the first time that the imposition of MIL-H-46855, a human factor engineering specification, did not mean that only the "dial and knob" boys would be involved. The HFE Personnel in several instances provided rationale to the design engineer in formulating his design. In co-authoring the Human Engineering Analyses Report a bond was formed that remained throughout the program.

What were the results? The IPPS design had a predecessor. The Predecessor which was not designed using Human Factor Engineering Personnel, required 1/2 a man to operate it during the mission and took several hours to "make ready" for mission. The IPPS was completely automated, essentially run by the pilot who had on his JIFDATS control panel, a three way switch (high resolution, off, low resolution) and two (2) fault indicators (BITE) representing the entire IPPS Sussystem.

In the case of the PRPV, it also had a predecessor. Its predecessor also was not designed with any emphasis on Human Factor Engineering concepts. This system requires almost two men to operate it during its mission and about the same to "make ready" for mission. The PRPV requires less than 1/2 a man during its mission and perhaps the same to "make ready" for mission. Figure 12 presents the control panel of the entire PRPV. The overall dimensions of this control panel is nineteen (19) inches by thirteen (13) inches. Controls in the PRPV's predecessor were spread amongst four (4) different consoles.

The cost for the Human Factor Engineering program based on MIL-H-46855 for the IPPS and PRPV was less than 2% of the total cost of the program. If the preceeding units were designed similarly, the "less than 2% cost" would have been returned to the user within the first year of operation.

From the discussion just completed it might seem that only a small portion of MIL-H-46855 was utilized, merely, Appendix paragraphs 30.3.1 to 30.3.9. This was not the case, emphasis was placed by the prime contractor on all subcontractors to systemize their approach to design early in the program. The requirements of these paragraphs forces this.

Other sections of this specification emphasize the other elements of the basic human factor engineering approach outlined in the beginning of this paper.

The requirement for mockups and models (paragraph 3.2.2.1.1) early in design proved extremely useful in making design changes and modification prior to actual fabrication and actually proved the impossibility of one approach in the design of the PRPV console.

The requirement for work environment and facility design (paragraph 3.2.2.3) not only provides for the anthropometric consideration of MIL-STD-1472 but also takes into consideration psychological aspects of human performance both in the normal and the emergency conditions.

The requirement of human factor engineering in development testing and evaluation (paragraph 3.2.4.1) is unique to a human factor engineering specification. It approaches an attempt to quantify the level of Human Factor Engineering involvement in design.

In summary, man-machine interfaces will continue to increase in complexity in tune with man's own evolution. Prudent use of Human Factor Engineering requirements such as those listed MIL-H-46855 can reduce these complex interfaces to child's play. In a period in which we are again trying to humanize the greatest system of all (LIFE), let's at least give the Human Factor specialist a

chance to show us the way.

References:

1. A manual of Work Study Technology, The Stanwick Corporation, Arlington, Virginia, 1970.

2. Engineering Psychology and Human Factors in Design, A. E. Javitz, ed., Electro-Technology, May 1961, P. 107.

3. Human Factors: Engineers Blind Spot, D. Meister and D. J. Sullivan, Electro-Technology, August 1968, P. 39.

4. Assigning a Value to Human Reliability, Herman L. Williams, Machine Design, July 1968.

5. Trends in Human Factors Research for Military Information Systems, Sylvia R. Mayers, Human Factors, 1970, P. 177.

6. Human Engineering Requirements for Military Systems, Equipment and Facilities, MIL-H-46855, February 1968.

7. Human Factors Engineering Systems Analysis Data, MIL-H-81444 (AS), September 1966

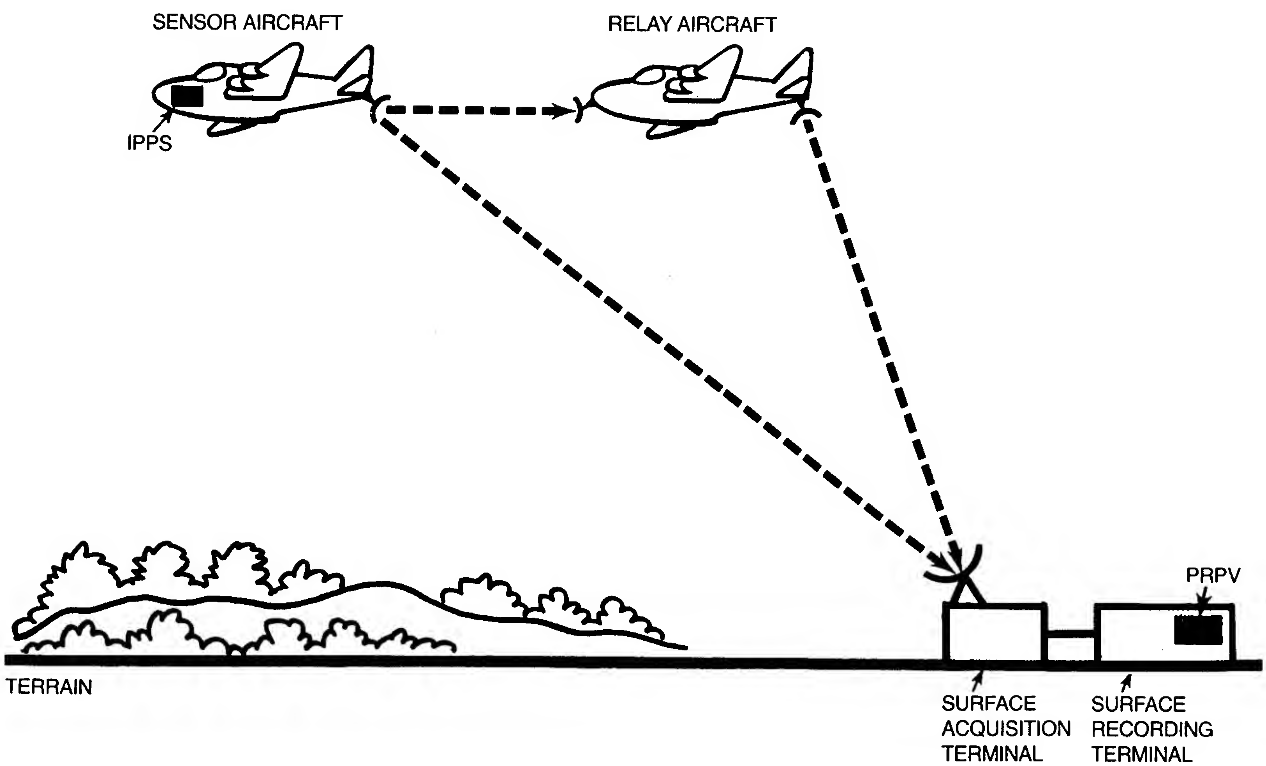
8. Human Engineering Design Criteria for Military Systems, Equipment and Facilities, MIL-STD-1472, DOD, October 1967.

9. Human Engineering Design Criteria for Aerospace Systems and Equipment, MIL-STD-803, U. S. Air Force, 1964.

10. Human Factors Engineering, 3rd ed., E. J. McCormick, McGraw Hill Book Company.

Figure No. 1 – Optical Illusion

Figure No. 2—Position of IPPS & PRPV in Jifdats Concept



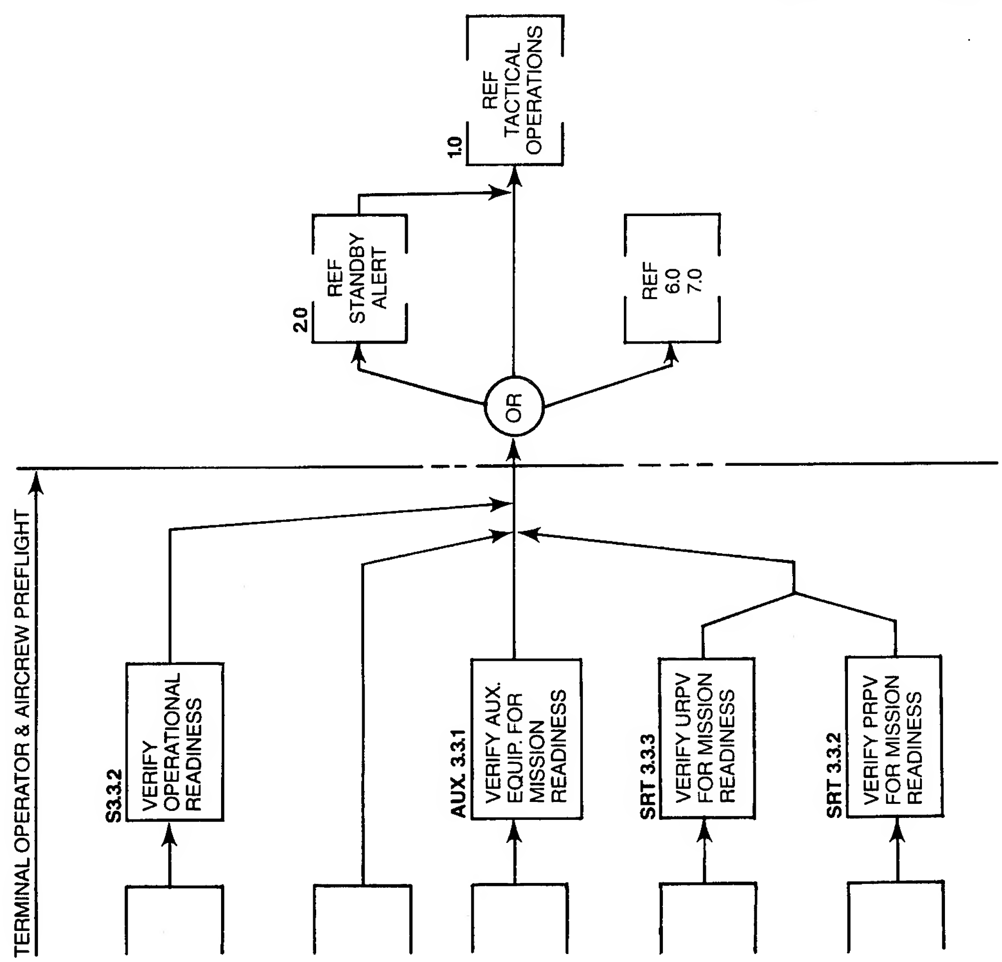
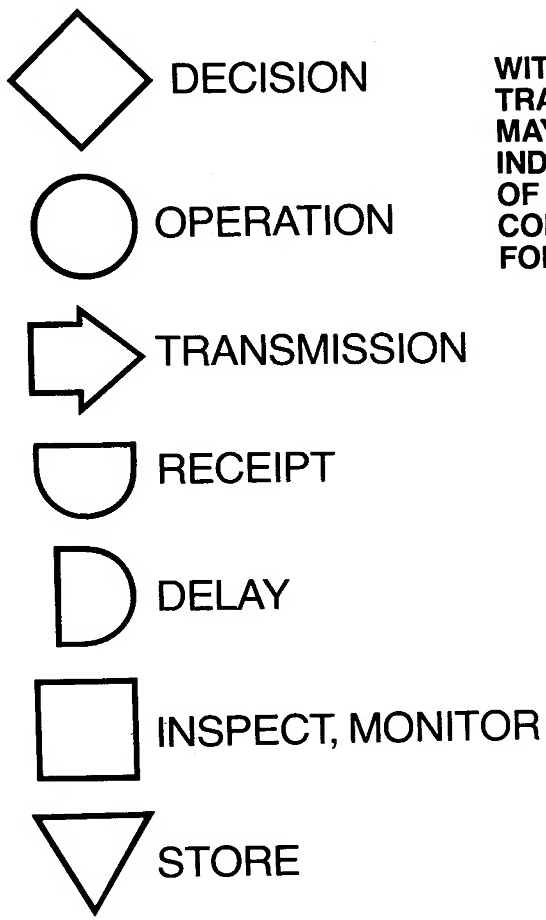


Figure No. 3—Portion of 2nd Level Functional Flow Diagram, Jifdats Secondary Mission

Figure No. 4—Mil-H-46855 Symbology

SYMBOLS



WITHIN THE SYMBOLS FOR TRANSMISSION & RECEIPT MAY BE FOUND ADDITIONAL INDICATIONS OF THE TYPES OF LINKS INVOLVED IN THAT COMMUNICATION AS FOLLOWS:

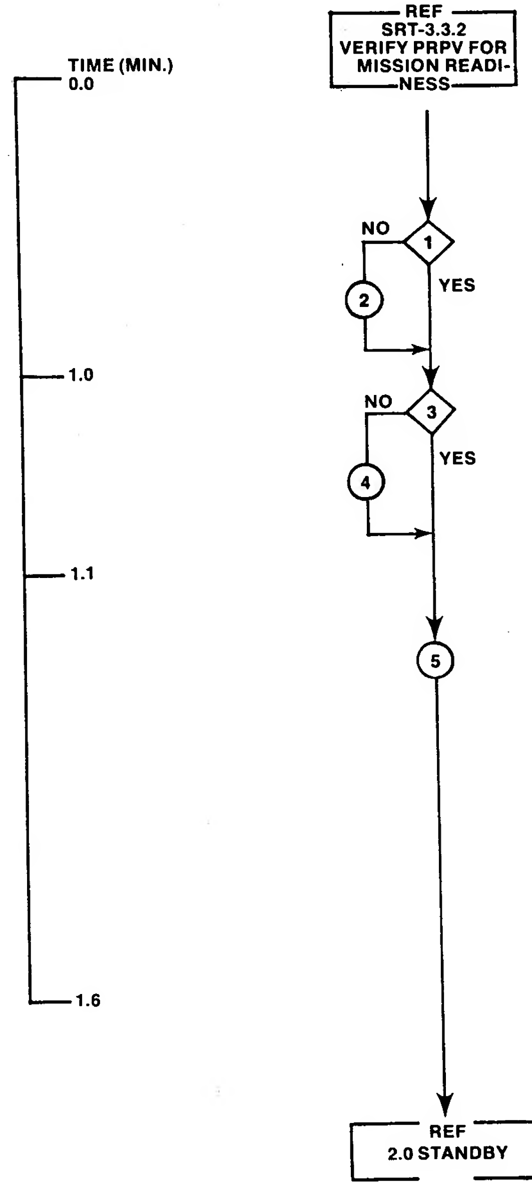
M—MECHANICAL OR MANUAL

E—ELECTRICAL

V—VISUAL

S—SOUND

Figure No. 5—Function & Time Analysis



FUNCTION NO.	FUNCTION
3.3.2.1	TEST RUN RESULT IN SATISFACTORY FILM IMAGE?
3.3.2.2	TAKE CORRECTIVE ACTION ESTABLISH SATISFACTORY IMAGES
3.3.2.3	PRPV IN "GO" CONDITION?
3.3.2.4	TAKE APPROPRIATE ACTION(S) TO CORRECT ANY BITE FAULT INDICATION(S)
3.3.2.5	NOTIFY OPERATIONS OF PRPV READY STATUS

Figure No. 6 — Functions Allocation Table

MISSION SEGMENT: STANDBY ALERT
FUNCTIONS & TIME ANALYSIS REF: 2.0 STANDBY ALERT

SUBSYSTEM: PRPV

FUNCTION NO.	FUNCTION	EQUIPMENT	PERSONNEL	RATIONALE	CRITICAL PERSONNEL-FUNCTION JUSTIFICATION
2.					
1	RECEIVE "STANDBY ALERT" SIGNAL	RECORDER PROCESSOR	PRPV OPERATOR	ELECTRO-MECHANICAL ACTIONS REQUIRED TO ESTABLISH MODE ELECTRO-MECHANICAL ACTIONS REQUIRED TO ESTABLISH MODE NEEDED TO PREPARE FOR FURTHER ACTIONS	
2	MONITOR AND EVALUATE BITE AND OTHER STATUS INDICATORS		PRPV OPERATOR	INDUCTIVE REASONING REQUIRED; INFORMATION REQUIRED FOR SUBSEQUENT DECISIONS	
3	ANY MALFUNCTION INDICATED?		PRPV OPERATOR	OPERATOR POSSESSES INFORMATION BASED ON MONITORING FUNCTION; NEEDED TO GUIDE SUBSEQUENT DECISIONS	
4	COORDINATE PLANS FOR CORRECTIVE ACTION WITH OPERATIONS		PRPV OPERATOR	FLEXIBILITY OF PLANNING REQUIRED; VERBAL COMMUNICATION REQUIRED	
5	TAKE CORRECTIVE ACTION TO RESOLVE INDICATED MALFUNCTION		PRPV OPERATOR	FLEXIBILITY OF ACTION REQUIRED; COST AND SPACE LIMITATIONS	
6	NOTIFY OPERATIONS OF MISSION READINESS		PRPV OPERATOR	VERBAL COMMUNICATIONS MAY BE REQUIRED; OPERATOR HAS KNOWLEDGE OF HIS OWN STATE OF READINESS AND THAT OF HIS EQUIPMENT	

Figure No. 7 — Operational Sequence Diagram

MISSION SEGMENT: TERMINAL OPERATOR AND AIRCREW PREFLIGHT
FUNCTION REF: SRT-3.3.2 VERIFY PRPV FOR MISSION READINESS

OPERATOR POSITION: PRPV

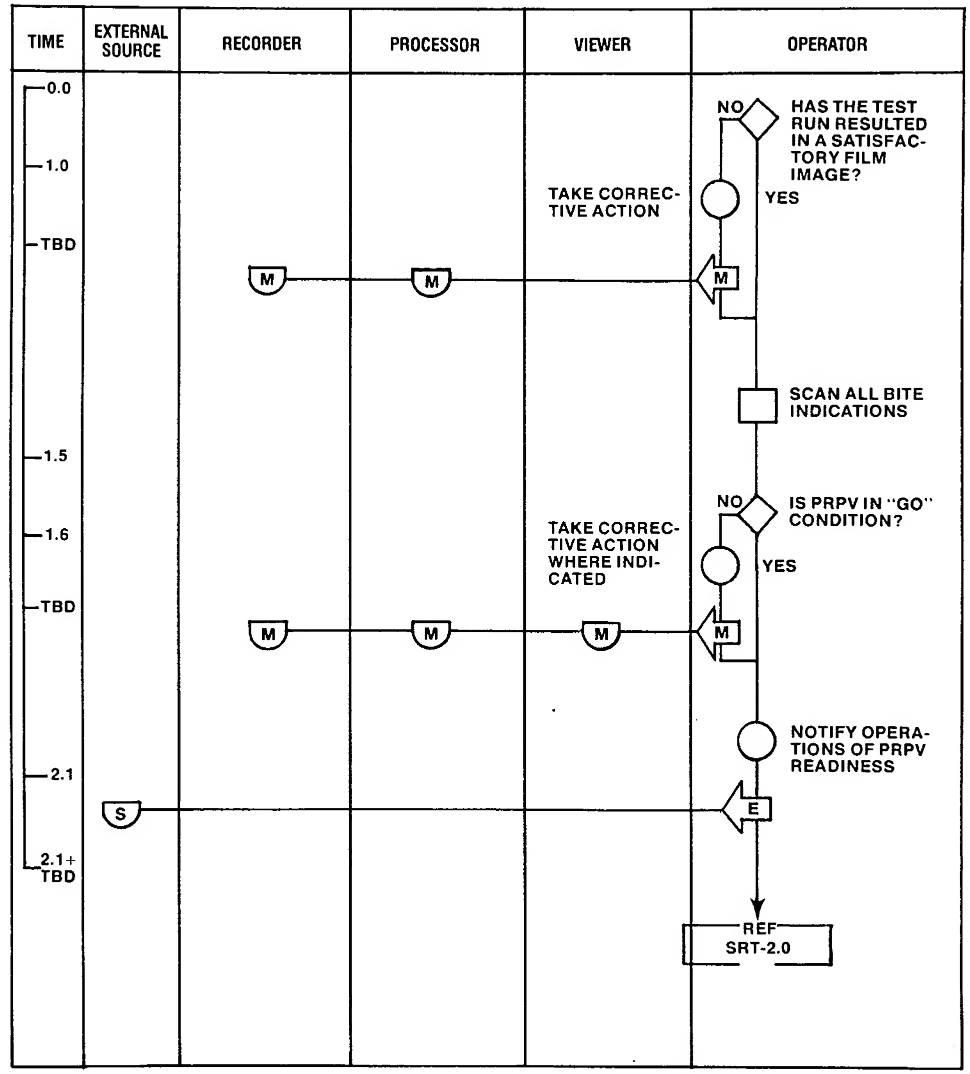


Figure No. 8 — Operator Information Requirements

MISSION SEGMENT: TERMINAL OPERATOR AND AIRCREW PREFLIGHT **OPERATOR POSITION:** PRPV
FUNCTION REF: SRT-3.3.2 VERIFY PRPV FOR MISSION READINESS

FUNCTION NO.	FUNCTION	SOURCE	INPUT REQUIREMENTS	PROCESSING	OUTPUT REQUIREMENTS	DESTINATION
SRT-3.3.2						
1	TEST RUN RESULT IN SATISFACTORY FILM IMAGE?	PRPV VIDEO TEST SIGNAL GENERATOR OSCIL-LO-SCOPE	(1) TEST FILM IMAGERY (2) PRPV PERFORMANCE SPECIFICATIONS (3) CRT DISPLAY OF TEST SIGNALS	COMPARISON OF TEST IMAGERY CHARACTERISTICS TO PERFORMANCE SPECIFICATIONS	(A) IDENTIFICATION OF FILM IMAGERY DEFICIENCIES	PRPV OPERATOR
2	TAKE CORRECTIVE ACTION TO ESTABLISH SATISFACTORY FILM IMAGE			CORRELATE UNSATISFACTORY IMAGE CHARACTERISTICS WITH LOCUS (1) FOR CORRECTIVE ADJUSTMENT	(B) MAKE ADJUSTMENTS TO APPROPRIATE COMPONENTS	RECORDER PROCESSOR
3	PRPV IN "GO" CONDITION?	BITE	(1) BITE INDICATIONS	PERCEIVE ANY UNSATISFACTORY BIT INDICATIONS (S)	(A) IDENTIFICATION OF COMPONENTS CAPABLE OF OPERATOR CALIBRATION	RECORDER PROCESSOR
4	TAKE ACTION TO CORRECT ANY UNSATISFACTORY CONDITION SHOWN BY BITE				(B) MAKE ADJUSTMENTS TO APPROPRIATE COMPONENTS (C) NOTIFY OPERATIONS OF A CONDITION THAT REQUIRES REMOVAL AND REPLACEMENT	OPERATIONS BASE MAINTENANCE
5	NOTIFY OPERATIONS OF PRPV "READY" STATUS	PRPV OPERATOR BITE	OPERATOR RECORD AND KNOWLEDGE OF PRPV CHECKOUT RESULTS AND PRESENT STATUS INDICATIONS	EVALUATE READINESS OF PRPV	NOTIFY OPERATIONS OF PRPV READY STATUS	OPERATIONS

Figure No. 9 — Control/Display & Communications Requirements

MISSION SEGMENT: TACTICAL OPERATIONS **OPERATOR POSITION:** SENSOR A/C SYSTEM OPERATOR
FUNCTION REF: 1.0

ITEM NO.	SPEC REQ'MT	INFORMATION DISPLAY	QUAL	QUAN	TOLERANCE	OUTPUT/CONTROL	DISCRETE	CONT.	COMMUNICATION REQUIREMENT	HUMAN ENG'RG CRITERIA
1	(75-2700-01) 3.4.1	INDICATE IPPS MODE FOR RESOLUTION AND TRANSPORT RATE 1. OFF 2. STANDBY 3. LOW RESOLUTION 4. MEDIUM RESOLUTION 5. HIGH RESOLUTION	X			SELECT IPPS MODE FOR RESOLUTION AND TRANSPORT RATE 1. OFF 2. STANDBY 3. LOW RESOLUTION 4. MEDIUM RESOLUTION 5. HIGH RESOLUTION	X			
2	(3.5.1.7)	6. DELAY	X			6. DELAY			ADVISE OPERATIONS AND SRT OF SCANNER STOP	
3	3.5.1.7	INDICATE PRIMARY SLACK BOX CAPACITY FULL/READY	X						ADVISE A/C COMMANDER THAT CAMERA IS INHIBITED	
4	3.5.1.6	INDICATE PROCESSOR BITE 1. GO 2. NO-GO	X						ADVISE OPERATIONS	
5	3.5.2.6	INDICATE SCANNER BITE 1. GO 2. NO-GO	X	X					ADVISE OPERATIONS	
		MISSION PLAN							RECEIVE NEW MISSION REQUIREMENTS	
6		CAMERA OPERATING	X							

Figure No. 10 — Operator/Maintainer Task Description

MISSION SEGMENT: FLIGHT LINE CHECKOUT
FUNCTION REF: SRT-3.2.5 CALIBRATE PRPV

OPERATOR POSITION: PRPV

TASK NO.	TIME, MIN & TENTHS	TASK	TOOL & AIDS	SPECIAL HAZARDS	CONSEQUENCES	DELAY EFFECT
1	CONT	MONITOR RECORDER STATUS INDICATORS				
2	0.1	DECIDE IF LASER BEAM INTENSITY IS SATISFACTORY				
3	1.0	ADJUST LASER BEAM INTENSITY*		LASER BEAM CONSTITUTES A HAZARD TO VISUAL ORGANS. BEWARE OF SPURIOUS REFLECTIONS OFF TOOLS INSERTED INTO PRPV	PERMANENT BLINDNESS OR OTHER EYE DAMAGE CAN RESULT IF LASER BEAM STRIKES EYE	
4	CONT	MONITOR FLUID TEMPERATURE				
5	0.1	DECIDE IF FLUID TEMPERATURE IS SATISFACTORY				
6	0.2	ADJUST TEMPERATURE TO REQUIRED LEVEL				
7	CONT	MONITOR SPINNER ROTATION				
8	0.1	DECIDE IF SPINNER ROTATION IS SATISFACTORY				
9	2.0	ADJUST SPINNER POSITION AND SPEED				
10	CONT	MONITOR FILM QUALITY*	MAGNIFIER		IF QUALITY IS NOT MAINTAINED, MUST WAIT FOR AIRBORNE FILM TO ARRIVE FOR PHOTO-INTERPRETATION; POSSIBILITY THAT SENSOR A/C MAY NOT RETURN	REDUCED TIMELINESS AND TACTICAL REACTION TIME MAY RESULT
11	0.1	DECIDE IF RESOLUTION IS SATISFACTORY				
12	0.8	ADVISE OPERATIONS OF DEGRADED IMAGE				
13	0.1	DECIDE IF IMAGE DENSITY AND CONTRAST ARE SATISFACTORY				
14	0.5	ADJUST CONTRAST AND DENSITY				
15	CONT	MONITOR CENTERING OF SWEEP				
16	0.1	DECIDE IF SWEEP CENTERING IS SATISFACTORY				
17	0.6	ADJUST SWEEP CENTERING				

*IDENTIFIES CRITICAL TASK

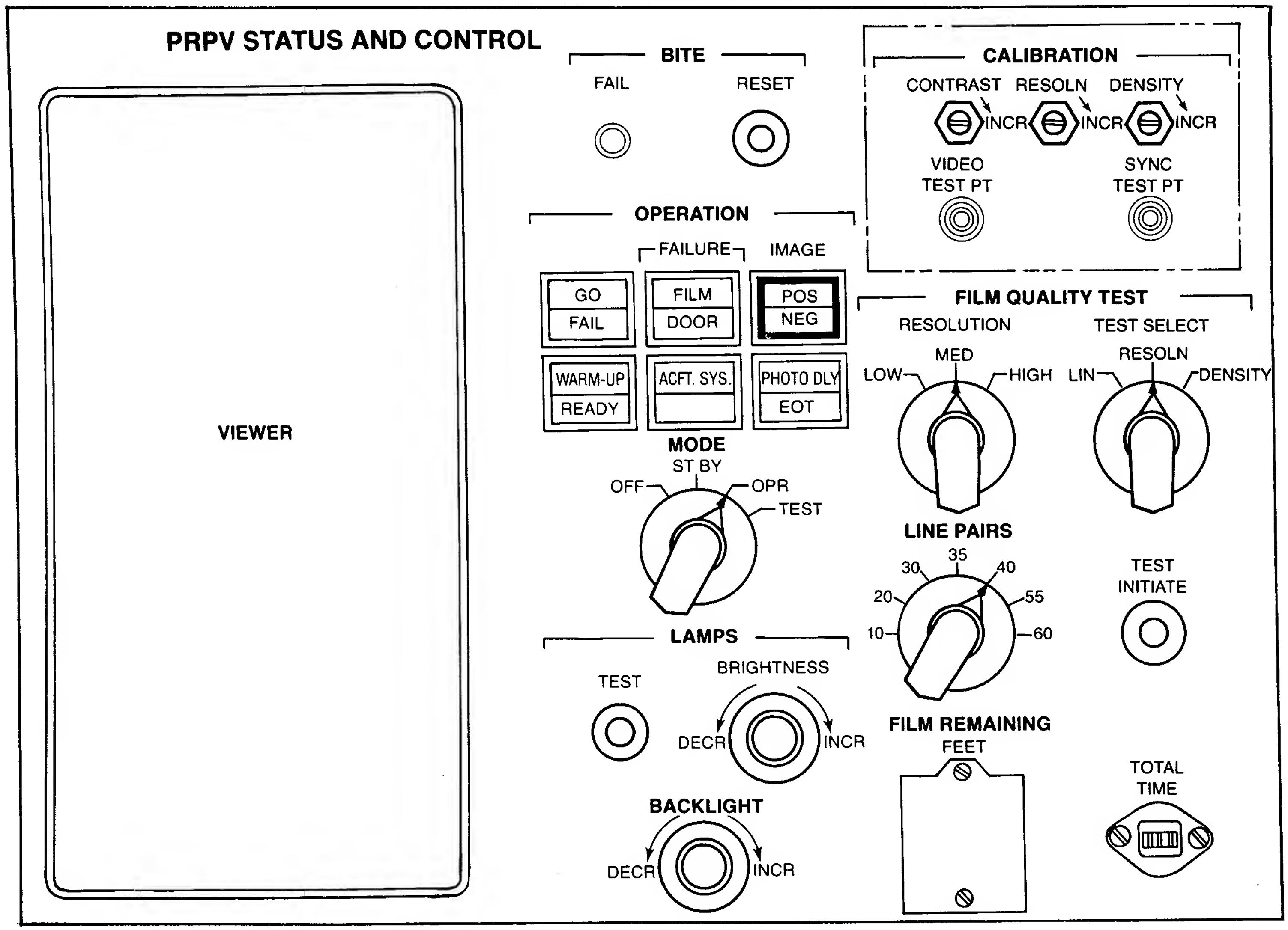
Figure No. 11 — Crew Loading Task Time Profile

MISSION PHASE: PRE-MISSION PREPARATION
FUNCTION REF: S-3.1 FLIGHTLINE SERVICE

OPERATOR POSITION: IPPS GROUND CREW

TASK NO.	TASK TITLE	TIME/MINUTES (0 5 10 15 20 25 30 35 40 45 50 55 60)	CRITICAL TASK JUSTIFICATION
S-3.1.1	UPLOAD FILM IPPS		
1	SPLICE CAMERA FILM LEADER TO PRIMARY SLACK BOX FILM LEADER		
2	INSPECT SPLICE		
3	SPLICE SATISFACTORY?		
4	INSPECT FILM ATTACHMENT TO TAKE-UP SPOOL		
5	ATTACHMENT SATISFACTORY?		
7	APPLY POWER		
8	ACTIVATE FILM TRANSPORT		
9	OBSERVE TRANSPORT OPERATION		
10	TRANSPORT OPERATION SATISFACTORY?		
12	HAS FILM SPLICE ARRIVED AT TAKE-UP SPOOL?		
13	STOP FILM TRANSPORT		
14	REMOVE POWER (TOT. TIME: 3.3 MIN.)		
S-3.1.2			
1	DISCONNECT FLUID LINES		
2	DISCONNECT TANKS		
3	REMOVE TANKS		
4	INSTALL FRESH TANKS		
5	SECURE TANKS		
6	CONNECT FLUID LINES		
7	INSPECT FLUID LINE CONNECTIONS		
8	CONNECTIONS FLUID TIGHT (TOT. TIME: 5.4 MIN.)		

Figure No. 12—PRPV Control Panel



DEVELOPING RELIABILITY IN A DEVELOPING COUNTRY

Julian Hilman
Israel Aircraft Industries, Ltd
Lod Airport, Israel

Summary

Developing Reliability in a new and developing country may appear similar to the process of introducing Reliability to an existing company in a developed country such as the United States, England or France. Some of the problems are indeed similar, but there are many new and unexpected problems associated with a developing country. In general, the problems of introducing Reliability are dependent upon the nature of the product, the labor force available, the management of the enterprise and upon government actions. Reliability is considered from the point of view of buying reliable products and also from the standpoint of producing reliable products. This paper describes specific problems based upon experience in a relatively advanced "developing country", i.e., Israel and particularly within its aircraft industry. Other problems are also discussed which may be of greater concern in other industries and in less developed countries. Information is given on resources required to introduce Reliability, a useful sequence for introducing Reliability tasks,and methods for demonstrating the effectiveness of Reliability activities in a developing country.

Introduction

A discussion of how to develop Reliability in a developing country should start with an understanding of two controversial terms, "Reliability" and developing country". In this discussion, Reliability shall be considered the practice of several, but not necessarily all, of the activities of Figure 1, as a separate function. in the development of an industrial product. Quality Assurance, Maintainability, Human Engineering and other related activities are specifically omitted from this list to simplify the discussion. The concept of Reliability <u>as a separate function</u> is used in this paper to show the conscious recognition of specific disciplines - although in actual practice, it is recognized, many of these "Reliability Techniques" may be considered "just good design practice" or even "only common sense".

When the question of "what is a developing country" was considered, it became apparent that all nations are "developing countries", but some have reached a more developed condition than others in one or more specific areas. In this paper, we are considering those countries which are just developing to the point where Reliability Technology is necessary for their continued development. Therefore, this discussion does not consider "what is a developing country" in the economic sense, but rather, under what conditions is any country (or industry or enterprise) ready to apply the techniques of Reliability.

Initiating Reliability

In the United States, it has been found that Reliability will be implemented only when someone needs it, demands it, pays for it and (hopefully) uses it. This may be the result of frequent failures of existing equipment or it may be impossible to design a complex system without using Reliability Techniques. In developing countries, Reliability Techniques may be introduced for the same reasons and also in three other ways: imitation, competition and option (see Figure 2.)

For a small, civilian aircraft, the customer does not usually demand Reliability as a specific requirement. However, when looking at how aircraft are designed and built throughout the world, it is almost impossible not to recognize the widespread use of Reliability Techniques (especially in America.) Even before Reliability is fully understood, there may be a feeling, in a developing country, that we "need to use some of those good techniques, too." This was the case in Israel when the Arava was designed. Reliability Techniques were introduced from the start and proved their worth repeatedly. In some countries (and some industries) the early acceptance of Reliability Techniques can be premature, while in other countries, acceptance may come only after a series of failures, customer complaints or disaster occurs.

The need for Reliability activities can also be a direct response to competition while trying to break into an established market for High Reliability products. This challenge has been recognized and met by a large sector of the Israeli electronics industry. However, for many engineers in a developing country, the initial confrontation with Reliability comes when they are faced with the selection of parts, components or assemblies and are offered various Reliability options along with various performance and price options. It is always difficult to decide how much Reliability is required and how much you can afford to spend for it - especially if you don't really understand what you are buying.

Buying Reliable Products

There are two aspects of Reliability to be considered:

1. as a user and buyer of reliable products

2. as a producer and seller of reliable products

As a user, the developing country must have a real and pressing need for the reliable system before asking for it. In addition to the usual considerations which come before Reliability (see Figure 3), there are also the considerations shown below.

Available Sources of Supply

Restrictions may exist for political as well as economic reasons. Economic restrictions may also include credit terms and balance of payments problems.

Reliability Comprehension

It is necessary to know what kind and how much Reliability to ask for in a specification. It is also necessary to know the possible consequences of requiring too much or too little Reliability in the specification.

Ability to Monitor or Measure the Reliability Received

This may be affected by the distances involved, the personnel and the test facilities available.

As a general rule, there are very few things which most developing countries must purchase for which there can be any justification for paying a premium for high Reliability. Such a requirement in a specification may be an indication that the equipment being purchased is more sophisticated than is really needed (i.e., the purchase requirement may demand more functional capability than is needed or may demand advanced and untried technology.) Very often, functionally simpler equipment, which may also be much less expensive, will be much more reliable. Some exceptions where high Reliability may be justified are shown in Figure 4.

Measuring Reliability

In the case of both the buyer and the producer, there is the problem of knowing how to measure Reliability - as a product characteristic or as a work effort (see Figure 5). The problem is somewhat reduced for small, inexpensive products which are made and purchased in large quantities. Then, standard techniques of testing may demonstrate a capability to withstand a sequence of environmental conditions and may even provide a modest estimate of the Reliability (in terms of failure rate or mean-time-between-failures or probability of successful operation.) Using existing standard national (or international) specifications such as U.S. Mil Specifications and Standards is a quick, simple, reasonably safe and reasonably inexpensive way to assure a minimum standard of Reliability. However, these specifications and standards require some understanding to assure that their use will result in the desired Reliability of the end product and will not be the cause of excessive costs.

For larger, more expensive items which are made and bought in smaller quantities, it may be necessary to look more closely at the intermediate results of the Reliability activities than at the completed hardware. The results of Reliability analyses can be reviewed for potential weak areas (or single point failures.) These potential problem areas can then be reviewed to determine the preventive measures that were taken. Design changes for Reliability improvement or special design features for improving Reliability can be identified. Reliability design criteria can be reviewed and exceptions can be identified. A history of failure occurences and the subsequent corrective actions can be of value in this review. All of these detailed reviews and investigations may provide a better "engineering confidence" than a Reliability number which is predicted with any "statistical confidence". In some cases, the detailed engineering review can provide the customer valuable insights concerning which of several available options is most desirable for his expected usage.

Developing Reliable Products

Realistic Reliability Goals (see Figure 6)

Although most engineers would like to make very reliable products, it would be well for every developing country to learn the basic business "facts-of-life" in the same way that every ambitious, new engineering graduate must learn them. It is important to understand the market at which the nation (or industry) is aiming and to have an appreciation of the level of Reliability required. It is possible to lose money as quickly selling premium products in a cost-conscious market, as by selling inexpensive "copies" in a sophisticated market. Many countries (and companies) which have started as manufacturing companies, using licensing agreements, believe that they can "move-up" to the high priced market by having the product look like the original, but is cheaper. It is not that easy. The labor force, tooling, investment and engineering effort to make a high Quality/Reliability product is much different than is needed to make a copy of a modest level Quality/Reliability product. The amount and type of test equipment can be more expensive than the original investment for the basic manufacturing facility. In some countries, this problem can be alleviated by government assistance in setting up a single facility to service the entire industry.[1]

Export vs Domestic Product Reliability

When considering the question of how much Reliability is required, for some products there must be a recognition of the possible difference in Reliability required for exported products and the Reliability required for domestic use.[2] This will not apply for products in which safety is a major concern such as aircraft. Where it does apply, it must be recognized that most developing countries with fledgling industry utilize protectionist techniques to encourage these young industries. At the same time, these young industries try to sell their products abroad.

Reliability is thus forced upward for export products, to meet the competition, whereas the domestic product need not utilize these Reliability techniques. While these protectionist tactics may be necessary for some products, this protection tends to become "a way of life" and can inhibit efforts to improve the product in the absence of competition. It is very easy to suggest that new industries should be set up in pairs or groups to encourage competition, but this is extremely difficult for a developing country which can barely raise capital for one company and which has the ability to absorb no more than the output of one company. The difficulty becomes even greater when the developing country has a "controlled economy", i.e., oriented to having many new industries financed by the government and having major operating decisions controlled by the government. It is difficult to justify setting up two or more government subsidized companies in competition for a market which may be too small to sustain one company. A solution which is being used in Israel, and possibly other countries, is to have the government finance one company and provide incentives for them to sell their products abroad. Then, when the company has become self-sufficient and their feasability has been demonstrated, the government looks for private investors who buy into the company and/or will set up additional, competing companies. This develops a "mixed" economy (private and government sponsored ownership) which allows the government to lead the private sector into activities which are beneficial to the country as a whole. The benefits sought are usually to improve the balance of foreign trade, maintain low prices within the country, help make the nation self-sufficient in critical areas and to promote the general economy. This can often improve the Quality and the Reliability of the products produced, as a by-product.

Technological Base Required

Countries like Japan and Israel did not build their reputations for high Quality/Reliability products overnight. Japan had a serious problem in overcoming a reputation for making inexpensive copies. This has been overcome and many Japanese products are the leaders for Quality and Reliability in the entire world. But the technological base which preceded these highly respectable products was developed gradually from working with lower cost products.

In Israel, the problem was different. Twenty-five years ago, there was little industry and the country was largely agricultural. The problems included a diversity of people (speaking 50 languages, from 70 countries - many of

them industrially backward countries) and the immediate problems of absorbing this population, feeding it and educating it. The problems were overcome by a number of wise decisions, some fortuitous circumstances and lots of hard work and belt tightening. Efforts were made immediately to provide an equal and high level of education for everyone. Schoolchildren were required to be fluent in a second language and were encouraged to learn a third. Foreign investment was encouraged, but controlled to certain kinds of products. Foreign expertise was imported and encouraged to stay as new immigrants. Vocational schools were opened and students were encouraged to enroll. Among the fortuitous circumstances were the existence of a fine technical university, the Technion, and also the famous Weizmann Institute. Both of these institutions have attracted many fine scientists and engineers of high calibre, have assisted scientifically based industry and have provided young Israelis with scientific inspiration and goals of the highest order. It was also fortunate that many of the immigrants had come from countries with a high technological state of development. These people had the skills to bridge the gap between the scientific and industrial leaders and those immigrants who were only minimally educated in modern scientific and industrial culture. Even so, it is only in the past five years that the Israeli reputation for Quality and Reliability in industrial products has begun to grow. A new generation had to come of age and industry had to develop before they could create highly reliable, new products with tightly controlled Quality.

Therefore, to develop highly reliable products, it is first necessary to be able to develop new products (with or without Reliability); it is necessary to have the capability to manufacture products with a high level of Quality; it is necessary to have sufficient capital, engineering talent, market research, technically competent labor and to have mature, aggressive and experienced management.

Industrial Differences (affecting Reliability)(see Figure 7)

Rapid Expansion. In addition to the special problems listed below, all the standard problems encountered anywhere when a small company tries to expand rapidly, are faced by every developing country in its industry.

Management. Management personnel usually come from academic backgrounds with excellent theoretical knowledge in specialized fields or from military and political backgrounds with strong motivation and skill at getting things done. However, in each case, the lack of experience in a large industrial enterprise also limits their appreciation of how and why products are built with inherent unreliability characteristics. The need for support organizations, clearly defined responsibilities and procedures are not always recognized, although they are mandatory and fundamental for assuring reliable products. However, these managers are usually quick to grasp new concepts and apply techniques which obviously produce results.

Motivation. Motivation of personnel is usually quite different than in the United States. Individual raises and overtime pay are less effective since salary levels are usually fixed on a national scale and high taxes limit any improvement in real income. Likewise, threats of layoff are not taken seriously in new and controlled-economy countries where job security is very firm and depends more on the state of the national economy than on the results of the individual's own immediate efforts. Basic discipline can be very different. In some countries, there is an automatic response to orders, whereas in other countries the order must appear reasonable and desirable (to the one receiving the order) before it is accepted and carried out. Motivation is much more successful when based on an appeal to national pride (probably in all new nations) and when there is public recognition accorded for personal achievements (particularly in small countries.) Supervisory personnel must set the pace/style for conscientious work attitudes. When people understand why their work is important, they quickly become involved and concerned. Although this is true in all countries, it is even more significant in new and developing countries.[4]

Recruiting. Obtaining experienced Reliability engineers in a developing country is a major problem. There is usually insufficient industry to simply advertise or to raid other companies for personnel. The sending of trainees to another country for a short course or bringing instructors from another country to give a short course are particularly ineffective. These courses are only an introduction to a few concepts and often prove the maxim that "a little learning is a dangerous thing." Training Reliability engineers in a university is, in my opinion, hopeless. Reliability must be learned on the job (supplemented by courses, if possible) and only after considerable experience has been attained in the areas of design or testing. Sending skilled engineers to another country for two to five years is highly desirable but depletes the home country of badly needed skills.

Another effective solution is to bring in experienced personnel on long term contracts. Such personnel must not only have extensive industrial experience in Reliability, Design and Test, but must be able to pass on their skills to local engineers and must be able to adapt their Reliability techniques to local conditions. Even with this plan, it requires considerable management support, freedom from administrative problems and available local talent. Unfortunately, there are no easy solutions.

Reliability in Israel

Government Activities

At this time, there is no coordinated government policy, nor specifications or guidelines for Reliability requirements in government contracts. Each procuring agency must decide for itself what Reliability requirements it needs. However, studies are currently being made which are expected to lead to standard requirements for some categories of government procurement. Where standard equipment is purchased in other countries, the existing requirements are usually applied, including Reliability requirements. When equipment is to be designed or manufactured within Israel, there is a wide variation in the Reliability requirements. When locally manufactured equipment is made under a foreign licensing agreement, there is often a question of whether the locally made equipment must be retested (for requalification). As more experience is acquired, it is becoming apparent that requalification is necessary and must be an accepted part of the cost of acquiring the local manufacturing capability. However, there is still much opposition to requalification on the grounds of added cost.

Education

There are many university level courses offered which relate to Reliability (such as Probability and Statistics, Strength of Materials and Stress Analysis), but few which cover the broad topic of Reliability. There are frequent short courses (one to two weeks) and seminars or symposia (one to two days) covering Reliability techniques and

tasks. These are usually sponsored by a university and a local professional society.[5]

Professional Societies

The meetings of the Israeli Section of the IEEE include special sessions devoted to Reliability.[6] There is currently an effort to activate an Israeli Branch of the Professional Group for Reliability. There is some debate on whether it is necessary to have a separate organization for Reliability or whether these needs can be satisfied within the ISQC (Israel Society for Quality Control) or the IEEE (general section) framework. This is particularly of concern when so many companies tend to combine their Reliability and Quality Assurance activities within the same organizational grouping.

The ISQC has been organized for less than two years, but already has 300 members and is planning an active program of technical meetings and seminars. There is consideration being given to eventual affiliation with the European Organization for Quality Control.

There are also technical symposia sponsored by the Engineers' Union which include new concepts and tutorial lectures. These symposia usually include a session devoted to Reliability.[7]

Industrial Practices

Industrial practices relating to Reliability vary widely, as may be expected when there is no central direction or incentives, as exist in the strong government policies in United States (for space and defense programs). As government needs for greater Reliability become formalized in contract requirements, there is a greater emphasis on this discipline in industry. Since the government requirements,which do exist, are usually in terms of results, computed or demonstrated, rather than organizational requirements or specific tasks, the organizations assigned responsibility for Reliability varies greatly. In those companies where the Design function and the Manufacturing function are closely related, the Reliability function is usually assigned to the Quality Assurance organization (or combined with it). Where the existing Design organization is somewhat independent, there is a tendency to assign the Reliability function to individual personnel within the Design group or to create a Reliability Group within the Design Engineering organization.

As a rule, the Reliability organizations are considered service groups (whose service is not always wanted). In some cases, they have the authority to review designs whether the designer requests this or not. Project leaders who have encountered difficulties in the field on previous projects or who have new demands thrust upon them by the customer are most eager to use improved Product Assurance techniques and to have guidance in designing for high Reliability. Unfortunately, this is still the exception.

Within the Israel Aircraft Company, there are a number of autonomous divisions. The Engineering Division is primarilly devoted to the design and prototype fabrication of aircraft. Within the Division, the Product Assurance Department includes both Reliability and Quality Assurance for the Division. The major activity relates to Reliability (which includes Safety and Maintainability). By comparison, the Manufacturing Division has a Quality Assurance Department which does not need to provide Reliability functions, as a rule. For specific problems which may arise, they will consult with the Product Assurance Department of the Engineering Division. In addition, they cooperate fully in joint activities such as failure reporting and corrective action. Again by comparison, the Electronics Division has its own Reliability and Quality Assurance Department. In this group, Reliability represents about one-third of their activity. They also cooperate with the Product Assurance Department of the Engineering Division, but are independent since they service the Design organization of the Electronics Division. Their relationship to the Product Assurance Department of the Engineering Division is closer to that of a subcontractor than that of a sister division. Similarly, each division has its own Reliability, Quality Assurance or Product Assurance organization tailored to its own specialized needs.

International Activities

Most of the Reliability data (such as failure rates) and techniques are based on literature and experience carry-over from the United States, including books, proceedings of symposia, U. S. Mil Specifications, etc. But, there is also use of standards, specifications and technical literature from many countries, particularly from Europe.

In addition, Israel is participating in the international data exchange program called "EXACT" (International Exchange of Authenticated Electronic Component Performance Data). In addition to its prime purpose of exchanging data internationally, the Israeli EXACT program is also serving as a focal point to pool test data from many companies in Israel who are vitally concerned with the Reliability of electronic and electro-mechanical parts and unbiased test results of these parts.

Conclusion

Developing Reliability in a developing country has most of the problems encountered in a more mature industrial society, plus many others due to local methods, customs, attitudes and working conditions. However, a developing country has not had time to develop fixed ideas on a new subject such as Reliability and there may be an opportunity for introducing the most effective techniques rather than merely standard techniques from other countries.

A developing country must be ready to enter the competitive area of new designs of reliable products, before it can afford a large investment in Reliability training and technology.(see Figure 8.) But if the industrial society of a developing country is to survive and flourish, it must eventually master and adapt Reliability techniques to fit both the local conditions and the demands of foreign competition.

References

1. Chen, T.T., "Q.C. Program of the Republic of China", Quality Progress Magazine, pp 17, May 1972

2. Armstrong, Willis C., "Quality and the International Economic Scene", Quality Progress, 17-19, July 1972

3. Kao, S.C., "Progress of Quality Movement in the Republic of China", Quality Progress, pp 14-16, May 1972

4. Editorial Staff, "Jishu Kanri Activities Activate Japanese Steel Industry", Quality Progress, pp 6-7, February 1972

5. Hilman, J., "Reliability as a Design Support Function", Reliability Seminar at Holon Institute of Technology, 10 June 1971, Holon, Israel

6. Hilman, J., "Boundary Conditions in Reliability Predictions", I.E.E.E. Symposium, 20 April 1971, Tel Aviv, Israel

7. Hilman, J. "Reliability for Avicnics Systems", Engineers' Union Symposium, 27 January 1972, Tel Aviv, Israel

F.M.E.A

DESIGN REVIEW

REDUNDANCY CRITERIA

FAILED HARDWARE ANALYSIS

RELIABILITY TEST PLANNING & EVALUATION

SPECIFICATIONS REVIEW

PARTS CONTROL

FAILURE DATA REPORTING

SUPPLIER RELIABILITY CONTROLS

RELIABILITY MODELS

RELIABILITY ALLOCATION & PREDICTION

FIGURE 1 : RELIABILITY TECHNIQUES

RELIABILITY WILL BE IMPLEMENTED WHEN IT IS WANTED, DEMANDED, PAID FOR & USED.

FAILURE PREVENTION

FAILURE CORRECTION

COMPETITION

IMITATION

OPTION

FIGURE 2: RELIABILITY INITIATION FACTORS

1. APPEARANCE

2. DELIVERY SCHEDULE

3. COST

4. PERFORMANCE

5. QUALITY

EXTRA CONSIDERATIONS FOR DEVELOPING COUNTRIES
AVAILABLE SOURCES
RELIABILITY KNOWLEDGE
ABILITY TO MEASURE RELIABILITY

6. RELIABILITY

FIGURE 3: PRIORITIES IN PROCUREMENT.

VITAL SERVICES

TELEPHONE EXCHANGES

RADIO COMMUNICATIONS

ELECTRICAL POWER

ENVIRONMENTAL HAZARDS –

EXTREME WEATHER

BIOLOGICAL PESTS

UNSKILLED PERSONNEL

EXPORT

PARTS, COMPONENTS, AND SUBASSEMBLIES

LARGE SYSTEMS.

FIGURE 4: HIGH RELIABILITY APPLICATIONS

REALISTIC RELIABILITY GOALS

EXPORT vs DOMESTIC PRODUCT RELIABILITY

TECHNOLOGICAL BASE REQUIRED

INDUSTRIAL DIFFERENCES AFFECTING RELIABILITY

FIGURE 7: INDUSTRIAL DIFFERENCES AFFECTING RELIABILITY.

TEST

WITH RELIABILITY ASSESSMENT

WITHOUT RELIABILITY ASSESSMENT

DESIGN RELIABILITY REVIEW

RELIABILITY ANALYSES

DESIGN CHANGES & RELIABILITY FEATURES

RELIABILITY DESIGN CRITERIA

(& EXCEPTIONS TAKEN)

FAILURE OCCURRENCES

(& CORRECTIVE ACTIONS)

FIGURE 5: MEASURING RELIABILITY

FIGURE 6: DEVELOPING RELIABLE PRODUCTS

RAPID EXPANSION

MANAGEMENT

MOTIVATION

RECRUITING

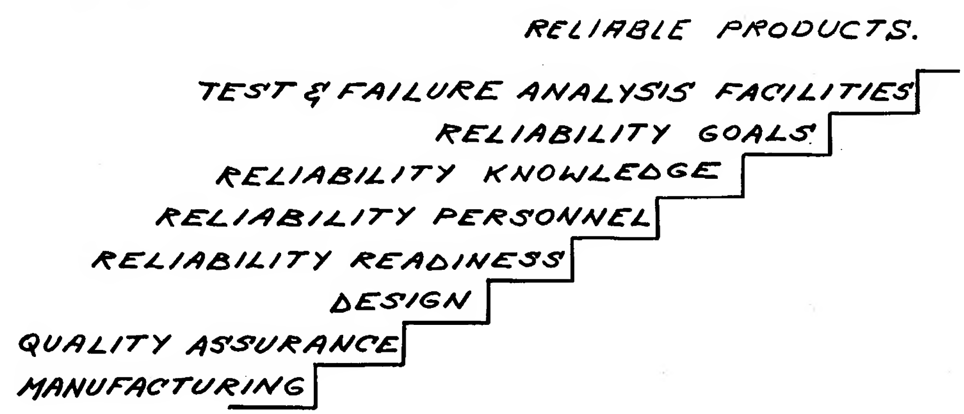


FIGURE 8 : STEPS TO RELIABLE PRODUCTS

SYSTEM EFFECTIVENESS
AND
THE ONE ERROR PER MAN PER DAY EXPECTATION.

SUMMARY.

Making reliability predictions at various stages of the development of a future equipment with new requirements is only a part of our mission. A reasonable prediction is the starting point and the reference for a more important task which is to tell our colleagues what to do in order to meet new challenges successfully and economically. A large part of our recommendations is directed towards the protection of all processes against human errors and this includes operations and maintenance.

Since it is no time to require special arrangements when a deadline and a fixed price have been accepted by the Company, we have to provide our management with good data at an early stage of negotiations with the buyers. When estimating we would rather have energetic pessimists sooner than sorry optimists later (or optimists who should be sorry). In many instances, in order to get a job, technical service or workshop managers underestimate costs by assuming that the new equipment will really work at its first test, that suppliers will not be late and that no reject will ever cause delay. This is a wrong attitude because only salesmen should be authorized to cut prices. Estimates in materials and in working hours should be realistic and make allowance for human errors.

In this paper we will deal with the impact of the human element on field reliability and on maintenance as an extrapolation of what has been learnt in the assembly workshop, we feel that these matters are also relevant to warranty costs and to customers' satisfaction.

INTRODUCTION.

Our subject has many sides, nevertheless it is basic to the future elaboration of a more consistent doctrine of reliability. The many sides of the human element in system effectiveness are philosophy, anatomy, physiology, psychology (no longer a part of philosophy for us) technology, probabilities and several more disciplines.

Naturally we do not expect anybody to master all disciplines pertinent to human behavior and a reliability engineer does not have to. He needs some simple data and he may have to be assisted by technicians conversant with processes he does not know well. With the human element as with hardware, a cautious engineer never releases a figure without specifying its conditions of validity and trying to get some feedback about his readers' interpretations. This is an anti-hara-kiri protection.

People produce potential failures in component fabrication, in equipment assembly, in system operation, in maintenance. Here we will limit ourselves to wrong actions, errors and mistakes, made within periods of acceptable performance of the worker. We will give up trying to cover more intellectual and vicious subjects such as wrong decisions, failures in communicating, forgetfulness, neglect, passivity, euphoria and other consequences of mental, moral, durable (illness) or permanent human deficiencies. This reduction to physical unrational actions is very artificial but it seems necessary as a first approach. For instance we will omit the story of the line worker who has a wonderful idea and humbly tries it with a disastrous result.

In our opinion, nationality or education have no direct bearing on the minimum error rate but they have some effect on practical results observed in a workshop or in the field : absent-mindedness, lack of discipline, assumption of smartness beyond working correctly may be linked to geographical, cultural, educational or age factors. We have been able to compare employees doing the same thing in various countries, differences in results are significant. But such differences do not affect the general bearing of our statements.

We will start with human errors on the assembly line because this is where we have the easiest access, then we will make some extrapolations.

With the one error per man per day as a minimum average we have a simple way of accounting for the human element when making predictions. We also expect to use this information in the study of the reliability of mechanical devices whenever we know how they are made.

HUMAN ERRORS ON THE ASSEMBLY LINE.

Historically it is our first contact with the subject of this paper ; about ten years ago a manager in charge of quality control at Company level discovered two ideas about human factors. His facts had been gathered in the electronic assembly workshop of six plants, some plants were producing radars other shops were producing communication equipment (already at that time results in missile and satellite production were more favourable but costs were correspondingly higher).

1st. idea. A worker makes at least one error a day on an average.

Day = eight hours.

We refer to qualified workers, knowing what they have to do.

In fact this result applies to good workers, very few make a smaller error rate. Workers with significantly more than one error a day, three errors for instance, were known as newcomers to the electronic profession or as careless or ungifted employees.

Defects are a way of evaluating the effectiveness of our training program.

2nd idea. A visual inspection detects 90% of the visible discrepancies which could cause a failure.

The more defects there are, the more of them go undetected.

A double inspection leaves about 1 % of visible defects.

These results apply to good qualified inspectors. Unseen defects remain as potential failures.

Example of a calculation based on these two ideas.

The most frequent defects are bad soldered joints, damaged components or the risk of a short-circuit. In failure rate documentation these defects are charged to the quantity of soldered joints, which in fact contribute to the largest share of the assembly error failures.

In real life some failures are suffered in the equipment test laboratory but we should not be dependent on that and we can discount equipment testing as a means of stopping assembly defects.

In eight hours a worker makes 400 soldered joints with his soldering iron (discrete component technology).

According to the first idea above, one of these 400 joints on the average is a potential failure and only 10% of the defective joints remain after inspection. After the delivery of the equipment by the inspection section the radio of bad soldered joints is:

$$0.25 \times 10^{-3}$$

If the equipment life is 10^5 hours or twelve years, we arrive at a failure rate of :

$$2.5 \times 10^{-9} \text{ failures per hour.}$$

This result obtained without the support of field reports and statistics is reasonable compared with data from the failure rate litterature, the RADC reliability Notebook gives 4×10^{-9} per hour as a high quality grade failure rate.

The above calculation is based on the average minimum human error rate, the practical failure rate must be based on the practical human error rate which can be known after a small (three or four) number of weeks of data collection for any particular assembly workshop. Tin wave soldering has improved joint reliability but human errors appear wherever they have a chance.

HYBRID AND LARGE. INTEGRATED CIRCUITS.

In the literature (1) we have found a way of predicting the reliability of LSI, and MSI based on the opportunities for human errors which can be found throughout the manufacturing process, adequate design being assumed. If the process is stabilized, the one error per man per day minimum average can guide us if we know the process, the duration of each manual operation, the task criticality, the quantity of potential failures that each kind of error can create and also the manufacturer's Quality Control. All such element should be known even before we order LSI's or even start detailed definition work with any manufacturer.

Although there are charts for failure rates corresponding to each step of production of a LSI, it may happen that we have no data about an essential, new operation. Then the first idea of the previous chapter is applicable. We have also verified that it works well with the assembly of transistors and standard integrated circuit in hybrid circuits.

Not only do prediction techniques based on the human element yield reasonable predictions, they are also a tool in process evaluation and in comparison when a choice has to be made between several machines.

For instance the failure rate of a bond is 2×10^{-9} per hour, the time for compression, heating and cooling is slightly more than a minute and production is under 400 bonds per operator per day, but opportunity for causing a potential failure is offered to the operator only during a very small part of the process, i.e. during positioning (pressure and heat are automatically controlled) and during handling.

Therefore the above failure rate can be verified well. Also we were easily convinced that for a particular hybrid circuit, the failure rate could not be smaller than 10^{-7} per hour, assuming good reliability of all semi-conductor components.

The above failure rate for a bond is higher than failure rates obtained with automatic equipment on standard semi-conductors. LSI and hybrid circuits are specific items which are often produced in small lots.

Our predictions may be proved optimistic. We may have to allow for the fact that LSI or hybrid circuit production is not a pleasant job (discipline, stern supervision, repetitive work) and the turnover is high among operators. Then the practical error rate we should consider may be two or three errors per man per day. Then if our prediction methods are still too optimistic there is either a problem of technology or a problem of production know-how.

Normally we have to assume that the process is sound. Otherwise reliability is zero and the problem is for some body else ; technicians, method experts, etc... We may try to help but we are no longer involved as reliability people. We always repeat that reliability problems are tied to the best of the bad parts, which implies that technical and basic quality requirements have been met and that obvious defective units are always rejected (figure 1)

OPERATOR RESPONSIBILITY AND QUALITY CONTROL.

We only mention a well known point about Quality Control Inspectors do not make Quality, they recognize it. Quality is made by workers, supervisors, process documentation, operation preparation, handling and storage. (We limit ourselves to conformance and omit design). Bad products should ideally be eliminated by the responsible operator. Unfortunately the employee may be unable to really "see" the result of his work.

If one of the essential prerequisities is not met either in specific circuit production or in electronic equipment assembly we can expect a high quantity of rejects after inspection and a high failure rate. Human errors will be multiplied five or tenfold and the customer will receive too many potential failures.

When we visit a possible subcontractor's plant we do not ask well known questions such as : "Do you have inspectors ?", the answer is always "yes" because no one likes suicide. If we ask "Do you budget inspection directly for each important contract?" or "Show me inspection times on your scheduling boards" or "Do your foremen have secretaries ? "Then the plant manager may call our vice-president about our lack of tact, which means that we are doing what we are being paid for (A foreman without a secretary is always under blueprints and papers and does not supervise a thing). (What is not budgeted is not fully done). (If inspection time is squeezed .... every one knows) (4).

The one error per man per day applies to manual work whenever operators have no excuse for producing defects, and when conditions are so good that they are responsible for the conformance of their production (2).

ERRORS IN EQUIPMENT OPERATION.

Honestly we have not been to obtain enough data which would support our former statements but we have not given up hope for some foreseeable future.

Equipments should be tolerant of operators'errors and it is desirable that damage causing errors be signaled for both maintenance and training purposes. It is well known that there is no difference between the unimportant mistake and the one which causes 500 deaths, the second one is the result of the unfortunate coincidence of a human error and the possibility of a serious accident.

The best experience concerning error in equipment operation is held by railroad and bus companies and more recently by airlines.

In a plant it is difficult to obtain the kind of data we want, but from the field it is nearly impossible. The best data collection program we know does not provide us with anything about human errors.

It is thus impossible to prove that one error per man per day is a realistic rate and that it is more correct than ten or 0.1.

We can give an example which is a simplification of a real situation :

Three switches are on a board, switch I has to be actuated many times a day, switches 2 and 3 should not be actuated in normal operation of an equipment controlled by the switch-board and pressing switch number 3 destroys a specific part one tenth of the time. From the frequency of the orders of specific spares by the customer we can infer an evaluation of the operator's error rate. The accuracy of such an inference may be rather poor but the solution would be better than the total absence of information.

If spares requirements are interesting, analysis of defective assemblies returned from the field is very valuable and well filled report forms are always appreciated by the reliability group. Only few customers provide us with this kind of information. Reports about human errors would stimulate research on better control layout, better labeling and improved manuals (3).

ERRORS IN MAINTENANCE.

In this chapter we have a result of some interest although the situation we will describe cannot be blamed only on maintenance technician random errors. We have observed failures occuring with a complex system and we have made charts showing the times of failures for each subassembly of each equipment and we have found that 60 % failures happening to a subassembly are repeated within 10 % of the MTBF of this subassembly, conversely 40 % failures happening for the first time remain single facts.

This experience is not an isolated one but it is our first attempt in analysis of quantitative practical maintenance effectiveness because we receive more thorough failure reports about this system than we generally get.

Among other things, we consider that the maintenance is to be thoroughly reviewed when a situation such as the above occurs. Insufficient technical maturity, secondary damage caused by a failure and poor manuals may be suspected. Our opinion is that human errors have to take their share of the blame for duplicated failures.

Since access to some sites may be difficult for many reasons, in parallel with actions by the responsible department, we have made models which illustrate the effect or ineffective maintenance on reliability and we have tried theoretically to reconstruct what is going on, a classical technique of process identification.

We imagine a device which can assume three states :

State 1. The device operates and its failure rate is λ

State 2. The device has failed, the maintenance crew is working on it, with a probability α the device is not well repaired, with a probability $1-\alpha$ it is correctly repaired and its reliability is restored.

State 3. The device operates, it has not been well repaired and its failure rate is $\eta\lambda$ with n greater than + 1.

Conséquence : Lacking any knowledge about the true condition of the device after its most recent failure, the failure rate is equal to the mean value between that in state 1 and that in state 3. (Figure 2).

Whenever there is no good reason for unsuccessful repair and if we are confident about the reliability of the device, we may reconstruct actual facts if we introduce human errors. If a repair is performed in one hour and if the repairman works eight hours a day, the one error per man per day gives a ratio of unsuccessful repairs equal to 0.125. If an unsuccessful repair gives the device a failure rate ten times larger than normal, the average failure rate is more than twice the normal one.

In a more elaborate model, we assume that after some time such as 0, 15 the MTBF, the device can be considered as correctly repaired and has returned to state 1 (figure 3).

The above explanations show how human error consequences can be traced when we have very little information.

CONCLUSION.

The reduction of human errors is important to ourselves, to our customers and to their crews. We only mention the effect of human errors on spare parts logistics. This problem still has to be treated in a scientific way but high dividends can be expected for the efforts.

Many wonderful modern tools tolerance neither error nor mistake and the knowledge of the general presence of human errors is very important to everybody since progress will depend on everybody's willingness to contribute. We are aware that we have some facts but that we are still missing some information which would justify the generalization of what has been observed in our plants.

The conclusion should not be the end of an effort but the beginning of further action and investigation.

BIBLIOGRAPHY

(1) Proceedings Annual Symposium on Reliability. 1970. Pages 364 - 380. LSI Reliability Assessment and Prediction. H. A. Lauffenburger.

(2) Proceedings Annual Symposium on Reliability. 1968. Page 75 Manufacturing aspects of Reliability Control. B. E. Ricks.

(3) Proceedings Annual Symposium on Reliability. 1970. page 114. Man as a part of the design environment. J. W. Christopher.

(4) Proceedings Annual Symposium on Reliability. 1967. Page 114. The view from the bottom. J. R. Isken.

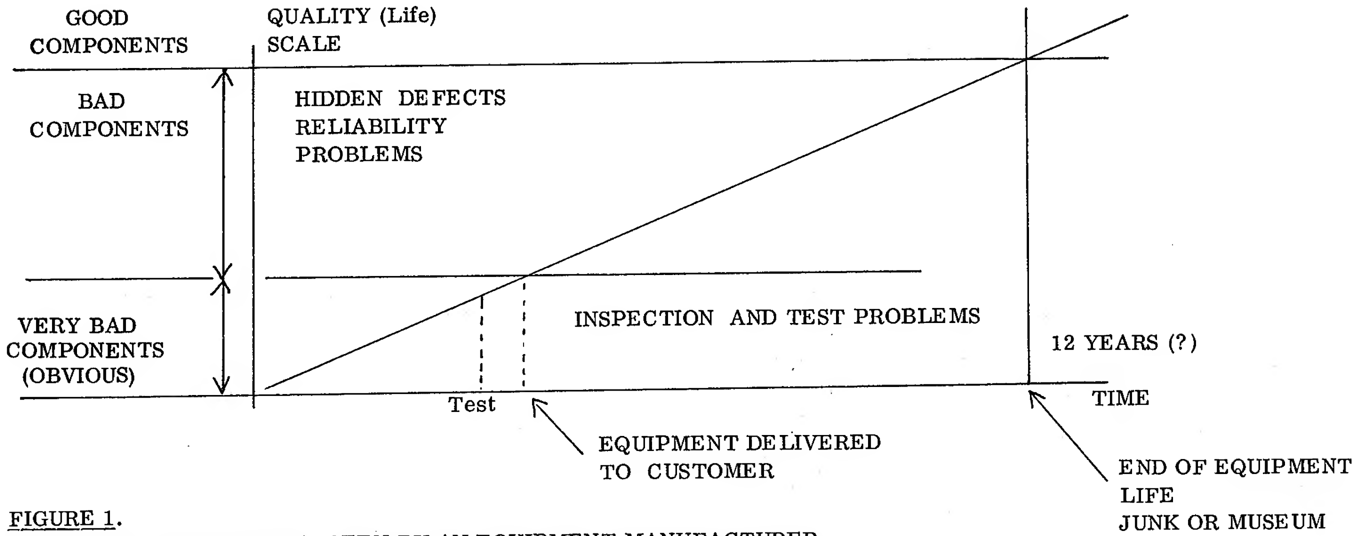


FIGURE 1.
COMPONENT RELIABILITY SEEN BY AN EQUIPMENT MANUFACTURER.

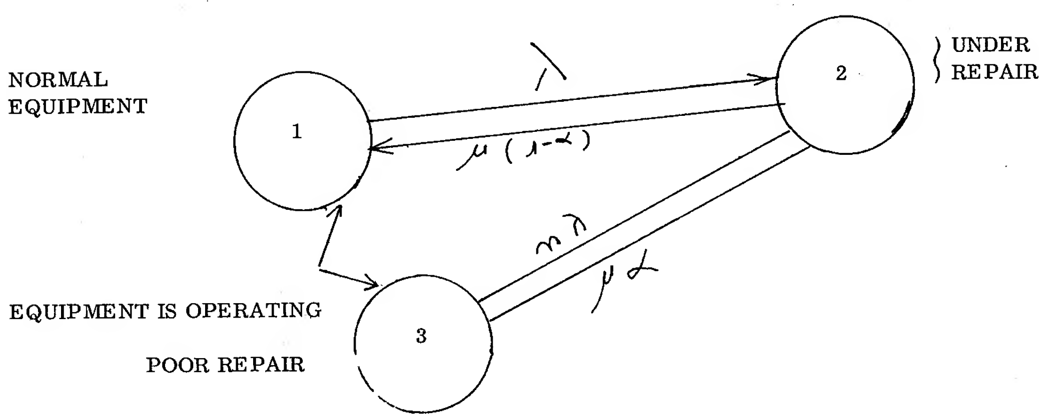


FIGURE 2.
RANDOMLY UNSUCCESSFUL MAINTENANCE.

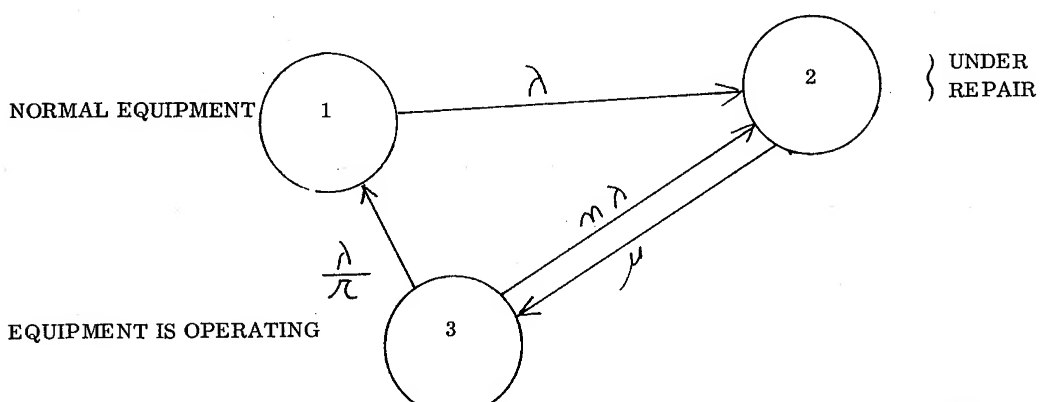


FIGURE 3. : SYSTEMATIC DOUBT.

STATE 3 : EQUIPMENT SUBJECT TO A DOUBT.

THE FUNCTION OF QUALITY/RELIABILITY ENGINEERING AND INTERNATIONAL TRADE

by R. M. Jacobs and J. Mihalasky
Newark College of Engineering
Newark, N.J.

Introduction

With the improvement in transportation technology, the world has "shrunk" to the point where products and services can easily be exchanged between nations. This exchange has been further spurred on by an increased interdependence between nations. The interdependence is the result of the nation's specialization in the manufacture of some, not all, products in order to maintain an efficient level of operation.

The results of this world "shrinkage", and the increase in specialization and interdependence, has led to an increase in international combines of countries. Witness the European combines such as the European Economic Community (E.E.C.) and the European Free Trade Association (E.F.T.A.) popularly known as the inner six and the outer seven, the European Committee for Standardization (C.E.N.) and the European Electrical Standards Coordinating Committee (C.E.N.E.L.). The United Kingdom, France and West Germany jointly established the Tripartite Committee to develop restrictive trade standards. There are other combinations such as the British Commonwealth group, the Pan American Standards Commission (COPANT), and combinations of African, Arab, and Socialist countries.

In dealing with these combines, the American exporting and importing companies have become familiar with the problems of tariffs, freight rates, import and export quotas, patent laws, and agency agreements. Now, however, these same companies will have to become additionally familiar with problems of product certification, standardization, and product liability and safety laws. This aspect of international trade has been largely ignored by American companies, and unknown to the general public because in the past U.S. standards were usually accepted and used by the other nations of the world. This was so because of our recognized superior technological position. But, times have changed, and the U.S. is no longer in that enviable position of having its standards treated as preferred standards. Both the developed, and the developing, countries of the world have embarked on crash standardization programs, followed by certification and product liability and safety laws.[1]

Initially standards were used as a device to control trade and maintain trade positions. Then it was found that standardization was essential to efficient production and marketing. In modern times, both the developed, and developing nations, have found that standardization serves to ease international trade, and promote the acceptance of each other's product.

Standardization as a Trade Barrier

Of late, standardization has again become a method of trade restraint and trade control. Blocs of nations have begun to develop standards that favor the members of the bloc, and put the nations outside of the bloc at a disadvantage. Additional barriers are created by requiring by law, certification, or conformity to these standards.

In 1970, the Tripartite Committee expanded its horizons to accept each country's certification on products, with membership open to all EEC and EFTA members. To date, the United States, as well as other non European countries, have not been permitted to participate in either CENEL or CEN. The result of this exclusion is that U.S. components, aside from possibly being more expensive, are even more difficult to market in Europe because they are subject to additional testing. The situation is further complicated for U.S. export companies by the fact that the Tripartite Agreement calls for inspection of manufacturing facilities as well as product testing requirements.[2]

It was widely believed by U.S. Industry that the above exclusion was a direct attempt to stem the flow of U.S. products into Europe. This means that a company exporting into CENEL countries will have to submit its products to third party testing, certification and inspection. This increases costs and is designed to put non-CENEL produced products at a disadvantage.

The International Electrotechnical Commission (I.E.C.) recognized the problem and has, after considerable debate and study, asked interested countries to participate in a management committee charged with the development of operating procedures, and rules for operation and financing. Participation by the U.S. is dependent upon financial support for the management committee& members.

Quality Marks

In addition, many nations have developed a system of standard quality marks. The quality mark is supposed to assure the consumer that the product bearing the mark meets certain quality, reliability, safety, and interchangeability standards. For example, the Australian mark is "AS, and has been administered, since 1955, by the Standards

Association of Australia. No foreign marks are permitted in Australia, nor is the AS mark registered in any other country. Foreign manufacturers are permitted to use the AS mark if they meet the criteria set forth by the Standards Association of Australia. An Australian consumer, given the choice, would probably pick a product with the AS mark over one without the AS mark. The consumer, as well as the Australian Standards Association, has the right to take action against any AS mark licensee.[3]

Canada registered in 1946 a "CSA" mark, in 1938 France started to use an "NF" mark, Germany was one of the first to use the "DIN" mark in 1920, India authorized in 1952 the "ISI" and in Japan two marks were authorized in 1949 and 1950, "JAS" and "JIS". Other countries such as the Philippines, the Union of South Africa, Great Britain, and the U.S.S.R. also employ quality marks. The United States, on the other hand, does not have a legally recognized national quality mark system, although the American National Standards Institute (ANSI) has under consideration, a project for such a mark which will bear the monogram "C". Use of the mark will not be compulsory but license will be required if such use is taken advantage of. Foreign manufacturers will be able to obtain permission to use the mark and foreign countries may also register their own marks. The U.S. is hoping to license the "C" mark abroad as well. As is the case in most developed nations, misuse of the mark can result in withdrawal or other legal proceedings.[4]

Current U.S. Status

It can be seen that U.S. leadership in the field of international commerce has and is experiencing a great deal of opposition from just about every nation in the free world and especially those countries which comprise the now ten-member European Economic Community. To complicate the situation, the U.S. is doing very little to keep pace with the rest of the world and this blasé attitude is serving as a catalyst to accelerate the movement. When one considers the fact that small nations such as Israel, Hungary and Ghana have been employing certification marks for years and the United States is currently "considering" the adoption of such a mark, one gets the feeling that our government has rested upon its laurels long enough. Further evidence of the American propensity not to comply with world opinion is the failure of the U.S. to metricate. The United Kingdom recognized this need for conformity several years ago and had the courage to initiate a program which it knew would be costly and confusing, but also, necessary.

To reverse these negative trends, it would appear that Washington had better enact some rapid, drastic legislation of a nature that will result in the restoration of the image of the United States as a world leader before this image deteriorates to a point at which it is so tarnished that it is irreparable.

The prospects for House passage of the International Standards bill HR8111 are not encouraging at this writing (8/20/72) even though the Senate has passed the International Voluntary Standards Cooperation Act of 1972 (S. 1798). Subject to voluntary international use would be engineering and commodity standards for products, processes, procedures, conventions, test methods, and their physical, functional, and performance characteristics. The standards, as indicated, are voluntary; business, for example, would not be required to accept and use them. Consumers, manufacturers, suppliers, etc., would all have a hand in the responsibility of promoting the public interest. The idea behind the act was to provide for representation of U.S. interests in bringing about voluntary standards and making agreements with other countries to assure compliance; to promote international trade; and to improve the balance of trade and payments.

In addition, Metrication bills, S-2483, HR-12307 and 12555 are being considered. The Senate has held hearings and passed its bill. The House has scheduled them for late 1972.

In December 1971 Mr. Richard O. Simpson, then Deputy Assistant Secretary of Commerce, reported that "The U.S. is actively participating in an attempt to formulate a General Agreement on Tariffs and Trade (GATT) Code on Standards and Certification. This Code, if followed, should ensure that standards and certification will serve to foster, rather than inhibit, trade. As presently contemplated the Code would apply to the full range of industrial products, would deal with mandatory as well as voluntary standards, and would apply equally to health and safety and environmental standards.

This code would give increased importance to the activities of the private sector's initiatives in international private nontreaty bodies such as ISO and IEC."[5]

A hearing by the Office of the Special Representative for Trade Negotiations was held on July 26, 1972 (with respect to standards) which brought out a great diversity of viewpoints.

There was a consensus on the value of international harmonization of standards, if such harmonization didn't weaken existing U.S. standards. However, many expressed doubts concerning any international scheme through which certification by a laboratory in one country would be accepted in other countries. The Electronics Industries Association has taken the position that:

a) only the manufacturer can provide quality assurance or certification for each product which he manufacturers, and

b) the independent body which makes periodic tests and audits of the manufacturer's quality function should be characterized as a Quality Assessment Body, not a Quality Assurance Body.

One of the hearing examiners stated in part "The GATT Code should not concern itself in any way with the preparations of standards by voluntary standards bodies. Voluntary standards bodies should be left to themselves, until their activities demonstrably raise public policy issues, at which point they should be legislatively controlled at the national level via the antitrust laws, or through other national legislation."[6] It is thus difficult to be optimistic about the possibilities of the U.S. adopting a Voluntary Standards Act and formally acknowledge ANSI as the recognized national standards institute and the coordinator of the private voluntary standards system; internationally, ANSI already represents the United States in ISO and IEC, the two major international standards writing bodies.

International Standardizing Bodies

There are in existence, international bodies such as the International Organization for Standardization (I.S.O.) and its electrical counterpart the International Electrotechnical Commission (I.E.C.) whose members are working on the task to "harmonize" the various national standards and certification programs so that they do not become trade barriers.

Historical

The International Electrotechnical Commission (I.E.C.) was created in 1906 when world leaders recognized the need for international standardization to promote international trade. In 1946 the International Standards Organization (ISO) was founded to which the IEC is affiliated. The object of the ISO is "to promote the development of standards in the world with a view to facilitating the international exchange of goods and services and to developing mutual cooperation in intellectual scientific, technological, and economic activity."[7]

The ISO and the IEC represent countries having four-fifths of the world's population. Between them the two organizations have published nearly 1,000 Recommendations - which cover an ever-increasing field and which represent an impressive tribute to international cooperation in the sharing of technical knowledge. There are also about 1,000 Recommendations currently in draft on which the national member bodies are being consulted in order to achieve the greatest measure of agreement before publication.

The ISO and the IEC have established close relations with bodies working in broadly similar fields. Some of these are intergovernmental, notably the regional commissions and other organs of the United Nations (the ISO and the IEC enjoy consultative status with the UN Economic and Social Council). Both of these organizations have as their prime function the development and promotion of standards that are mutually acceptable throughout the world. To carry out this function a series of Committees, each having technical experts from member countries, contribute of their time and expertise, preparing procedures, standards, definitions and guides in the form of Recommendations. The representatives of National Standards organizations who participate in these deliberations are expected to aid in generating the international standards and encourage their adoption in their respective countries.

In many countries these standards have the status of government regulations to which importers, exporters and manufacturers must adhere. In other countries such as the United States the adoption of these standards are basically voluntary by the manufacturers and others involved. These standards, developed by volunteers, are adopted by state, municipal and school board authorities as part of building, operating, zoning and licensing codes. They thus become, in fact, law, but are enforceable in those jurisdictions only. It is the intent of U.S. standards experts and the American National Standards Institute representatives to develop compatible Standards and Evaluation procedures which are useable in both types of requirements, the voluntary and the directed or regulated. Voluntary standardization systems are not governed by government regulations or law but by the buyers of a product. Since the buyer usually defines the characteristics desired these become the product requirements. The buyer elects to buy a standard or a non standard item. The public has this choice also and it is axiomatic that if enough buyers fail to purchase the product, standard or not, it is usually taken off the market. Thus the buyer voluntarily creates or defeats a "Standard". Manufacturers are interested in selling products to produce a profit and are motivated by volume.

Mr. William McAdams, President of the U.S. National Committee of the I.E.C., on July 26, 1972, before the Office of the Special Representative for Trade Negotiations on Possible GATT Code of Conduct for Preventing Technical Barriers to Trade, stated in part ......"We (the U.S.N.C.-I.E.C.) have one other serious concern that we believe ought to be examined carefully by your office. The background information supplied (on GATT) notes that more than 800 non tariff barriers (NTB's) restrict international trade at present, but does not attempt to list them. It is our impression that many of these 800 affect United States exports more than they do our imports. We suspect that the different standards used in the United States, as compared to much of the rest of the developed world, may be one of the major barriers we present to the foreign manufacturer attempting to export to the U.S., whereas other nations have more--and more complex--NTBs, in addition to standards, that affect our exports to them."[8]

It is therefore painfully obvious that the U.S. is attempting to maintain a leadership position in world trade by participating in I.S.O., I.E.C., C.E.N.E.L. and G.A.T.T. activities. However, without financial and moral support from governmental agencies and private industry these efforts by a few

volunteers may not succeed.

Foreign Product Liability Laws

Following close after the standardization and certification programs, are the product liability laws. Consumer protection organizations from over 30 countries have united into a group called the International Organization of Consumer's Unions. They have united on an international level to urge, among other things, a strengthening of consumer protection laws. Within the last several years a new product liability law has been enacted in Germany, which will probably become the model for the rest of Western Europe. This law applies to both domestic and foreign companies, protection for workers or operators of equipment from injury by defective product. The law also protects third parties. In some cases the manufacturer now has the burden of proving the cause of the experienced defect, and the manufacturer may be liable if he could foresee that his product could be unsafe even though he could not foresee the specific injury in question.[9]

This compares with the recently enacted Occupational Safety and Health Administration (OSHA) and the older doctrine of strict Liability contained in the Restatement of Torts (Second) Section 402A.[10]

Other European legislative acts which have been enacted recently are the British Trade Description Act, which came into force on November 30, 1968. The function of the Act is to deter misdescription, and to discourage its repetition, by the threat or fact of criminal proceedings. These could lead to fines, in serious cases unlimited in amount, or to imprisonment, or both. The Act does not require that a description be given, but where one is given it must be the truth. In the last year over 2800 complaints were handled by local enforcement personnel who have the ability to fine and imprison for violations.[11]

In Belgium new legislation, passed in 1971, now requires the processing of consumer complaints, which in the past took months or even years, to be handled within eight days by sellers of product.[12]

In addition, new legislation requires clear price and weight marking, prohibits misleading advertising, the "knocking" of competitors, sales at less than cost, 'pyramid selling', and the forcing of payment for unsolicited goods. It also controls marking in sales.

Russian organization is such that a factory can develop a new design and new process and receive credit for its ideas. The factory, however, cannot employ the design, process or product until the cognizant Ministry approves the plans. When approval is given, all factories in the Soviet sphere of influence receive copies of literature, drawings and test reports. By an established date, all producers of that product must follow the State procedures.

Thus, the Russians have nationwide standards for technological processes and product so that a piece of furniture purchased in Moscow is the same piece of furniture in Kiev or any other city in the Soviet sphere even though a different factory produced it. This effectively eliminates product shipment from one end of the country to the other. It eliminates competition since all companies are State owned and have their own marketing areas. But at the same time, it assures that the product has been tested as thoroughly as the Ministry decides is necessary. It is unlike anything we have tried in the United States.

Quality - Reliability Responsibility

One of the key underlying principles of all the items discussed above involves the measurement of the adequacy of a product to satisfy a set of standards or specifications, whether for certification, acceptable product listings, import or export licenses. Thus the Quality/Reliability Engineer is directly involved in every phase of standardization. Although he may not participate (he should) in the development of a requirement he is;

1. the Evaluator

2. the Compliance Certifier

3. the last manufacturer's representative to see the product before it leaves the fabricator's control.

4. the representative that testifies that;

    a. it was shipped in good condition
    b. it worked in tests
    c. it contained warnings and labels
    d. it was empty of combustibles
    e. other products like it were in operation.
    f. other products like it had not failed.
    g. there was no prior indication of failure, trouble or misuse.

5. the company representative who could be subject to criminal action if he was aware of defective conditions prior to the product's shipment.

Quality - Reliability International Involvement.

About ten years ago a number of specialists in the Quality-Reliability field recognized the need for international standardization of the many facets of Quality Control and Reliability terminology, techniques, formula, and procedures as it would eventually effect world trade. An organizing committee of these specialists met and petitioned the I.E.C. for Technical Committee status with the express desire to carry out the mandate of their concern. In 1965 the I.E.C. created Technical Committee 56 on Reliability of Electronic Equipment and Components used therein. Initially the American Society for Quality Control financially and morally supported the Secretariat for this Committee. Shortly thereafter A.S.Q.C. ran out of funds

and had to withdraw it's financial support.

However, the U.S.N.C. continues to be operated by volunteers with out-of-pocket expenses provided by a small number of Corporations. The American National Standards Institute (A.N.S.I.) and the Electronics Industries Association (E.I.A.) support the work by providing many routine administrative, office and mailing services.

During the General Meeting of the International Electrotechnical Commission in Washington, D.C. (May 1970), the United States Delegation of Technical Committee 56 - the committee on "Reliability of Electronic apparatus and parts used therein", proposed that the IEC study the possibility of a truly international Quality Certification scheme for Electronic Components. The U.S.A. Delegation also recommended that TC 56 be permitted to establish a Working Group to formulate the operational requirements and the appropriate protocol for this scheme. Both of these proposals were accepted by the IEC governing Council. Many reports and analyses were developed by participants in the special Working Group 7, which was headed by I.E.E.E. Fellow, Dr. Leon Podolsky, a Vice President of the I.E.C., U.S. National Committee. The I.E.C. Council, after a detailed study of the International Quality Certification plans and supporting reports, decided to accept the overall responsibility for the development of detailed procedures for Electronic Components as the initial commodity.

At the present time, the I.E.C. is establishing a management commitee to establish the rules and procedures of a Quality Certification Procedure. This committee will consist of two representatives from each country interested in the Certification Plan, each will be required to contribute a stipulated amount for the organizational and initial operational expenses of this group.

To fully implement the current plan, it will be necessary to establish in each country a National Supervisory Inspectorate. It's purpose will be to oversee, monitor and review the activities of the companies desiring participation in this International Quality Certification Procedure. It is expected that the United States members of I.E.C. TC-56 will participate in preparing some of the detail procedures of Quality Organization, Equipment Calibration Systems and Test Specifications.

Mr. A. Okun, Chief Delegate of the U.S. National Committee TC 56 recently stated, "If a workable model is established for Electronic Component Parts, it is my belief that a Quality Certification Procedure will be established for other products including home appliances, business machines and home entertainment equipment as well as other types of industrial equipment.[13] Trade associations such as the Association of Home Appliance Manufacturers, the Business Equipment Manufacturers Association and the National Electric Manufacturers Association should make themselves aware of the progress of this activity. The entire purpose of this international plan is to establish a means of accurately defining the quality as well as safety of products we, as producers, offer in international trade."

Various members of the U.S.N.C.-TC 56 represent professional trade organizations having an interest in the deliberations. These members then report to their respective groups for opinions, assistance, comments, etc. Frequently these organizations provide initial documents for consideration by the USNC-TC 56 and then by the IEC-TC 56 delegates.

<u>Technical Responsibilities of the Quality/ Reliability Specialist.</u>

The quality of a product is dependent upon each part of the process by which it is conceived, produced and brought to the customer. No amount of care and attention during manufacture can overcome deficiencies in design, while faulty material or purchased parts can sabotage all the skill and ingenuity of the cleverest designer. Even where a product itself is entirely satisfactory it can suffer from poor packaging, late delivery, faulty installation and inadequate operating instructions. Clearly, the full benefit of quality reliability assurance practices cannot be obtained unless the management control and philosophy is applied throughout the system. It must start with the first stages of design and extend through into the testing and operational field with proper arrangements for feed-back throughout.

While it is important to elaborate on the detail of every element of the system it is necessary to suggest that the reader refer to specialized reports on each subject. However, there are a few functions which appear to be more important to the subject of international trade than others in the Quality Reliability sphere.

<u>Standards</u>

The first is that of preparing standards and the Q/R Engineer is directly involved as is evident by the definitions put forth by Roy Trowbridge, President of ANSI.[14]

"An engineering standard has been defined as a technological practice described in a document to assure dimensional compatibility, quality and performance, uniformity of evaluation procedure, or uniformity of engineering language. It may, typically, prescribe screw-thread dimensions, clothing sizes, chemical composition and mechanical properties of steel, methods of test for sulfur in oil, or a code for highway signs."

Separate standards are generally issued for dimensional specifications, quality specifications, test methods, and descriptive practice. Obviously, the role of the measurement unit varies with the application.

Dimensional standards are needed to guarantee that a product or system will function, or to guarantee parts interchange-

ability.

Quality and performance standards assure (a) a quality level adequate for the required service, and (b) uniformity in quality from one item to another. These are dominant factors in safety standards. Additionally, standards may call out quality levels, Mean-Time-Between (or Before) Failure, maintainability and availability levels and the appropriate confidence requirements.

Test method standards provide a common basis to evaluate both materials and products. They establish standardized procedures to determine critical dimension or product quality, and are essential to determine compliance of a product with a specification.

Descriptive Standards include codes, symbols, sampling and other statistical terminology, format for engineering drawings, and other descriptive engineering practices.

The primary elements of standardization are:

1. Preparation of standard terminology.
2. Preparation of a statistical sampling procedure for judging conformance and compliance.
3. Preparation of the testing method and equipment, and the method of analyzing test results.
4. The writing of the specification to incorporate items 1 through 3.

Standard writing is an inherently technical process. In this country neither congress or the enforcement agencies have unilaterally assumed responsibility for writing standards. Most standards have been written by private voluntary standards groups. The enforcement groups have looked on from the sidelines, periodically expressing opinions.

Industrial laboratories support the specification writers with the testing and research associated with the preparation of a standard. These same laboratories generate data which is later used to update these standards. Modern mathematical techniques, and methods of prediction, augmented by the computer are also used to maintain the validity of standards.

<u>Certification</u>

The process of certification attests to the product's purchaser that the specification (standard) has been met. It attests to compliance with quality, reliability, and product safety levels by the product in question.

The certification process calls upon the quality control function to ensure this compliance. Here again, the laboratory, both in house and consultive, serves to provide the compliance data.

The document of certification also usually calls for compliance information such as test results, sampling results, and other quality level indicating paperwork.

<u>Product Liability</u>

The laws of more and more countries are reflecting the mood of the international consumer movement and incorporating stricter liability laws into their legal system. Therefore it is incumbent upon the manufacturer to protect himself from liability exposure. Some of the approaches to product liability exposure minimization include formal design review, which includes fault tree analysis and failure mode and effect analysis, data feedback and analysis systems, product testing, and reliability and safety prediction models.

In recent visits to the European countries discussing the effect of standardization on quality of product and the laws reflecting product quality, the authors found the following situation:

- Western European Countries (including Israel)

  In general Europe follows an approach which makes a personal injury a possible criminal situation. They also permit civil suits. The most celebrated of these in recent years is the Thalidomide babies, involving a German manufacturer. At the time of this writing another major situation is developing in France where some twenty children have died from the use of a powder called Bebe. Essentially they can and do hold individuals personally responsible for their actions - there doesn't appear to be immunity as we have it in the U.S., when the individual works for a Corporation.

- Eastern European Countries

  The law is essentially the same as Western Europe. However, since an injured person receives unlimited free medical attention, continues to receive full compensation as if he was still working and the product is replaced, there is no economic loss to recover. This essentially compares with the so-called No-Fault Insurance and Workmens Compensation practices.

The usual discussion for compensation for pain and suffering is not often raised as an issue, since the injured party would be suing his own government.

<u>The Quality/Reliability Engineers' Methods and Techniques</u>

The Quality/Reliability specialist is deeply involved with all of the items discussed above. He can and does contribute to the solution of international trade problems. He has become expert in one or more of the many facets usually associated with Quality Control and Reliability Engineering. When the detailed techniques are compared with the

statements made earlier about standardization, certification and product liability, it becomes easier to justify the statement that the quality/reliability engineer is in a commanding position to aid the business executive with some of his international trade problems.

REFERENCES

1. "New Product Liability Law" - German American Trade News, Jan-Feb. 1970 (p. 12-14, 21)

2. Rockwell, W.H. "International Trade: The Impact of Certification", The Magazine of Standards, July 1970, p77-79.

3. International Organization for Standardization, Index of Marks Indicating Conformity With Standards (Jan., 1971)

4. Rockwell, W.H., "The ANSI Certification Program", Sept. 13, 1971 American National Standards Institute.

5. Simpson, Richard O., "International Standards and Certification and it's Relationship to Trade", U.S. Department of Commerce News, Washington, D.C., Dec. 8, 1971.

6. Private report to the officers of the American Council of Independent Laboratories on the results of the GATT hearings.

7. Podolsky, Leon; "Name of the Game" American National Standards Institute, 1969.

8. McAdams, W.A.; "Statement before the Office of the Special Representative for Trade Negotiations" on Possible GATT Code of Conduct for Preventing Technical Barriers to Trade, June 26, 1972.

9. "New Product Liability Law" - German American Trade News Jan-Feb. 1970 (p. 12-14, 21)

10. American Law Institute, Restatement of Torts, 1965

11. Calascione, J.; "Let the Seller Beware" International Consumer, Vol. XIII No. 1 Spring 1972, The Hague, Netherlands. p. 12.

12. "News Page" International Consumer, The Hague, Netherlands, Vol. XIII No. 1, Spring 1972, p.19.

13. Okun, A.M., "Test & Inspection to Avoid Liability of Imports" Product Liability Prevention Conference Proceedings, Newark College of Engineering, Newark, N.J., August 23-25, 1972, p. 107-110

14. Trowbridge, R.P.; "Engineering Standards - What and Why?" Quality Management of Engineering, June 1972, p.32

15. Standardization, UNIDO Monograph on Industrial Development, No. 12, United Nations, New York, 1969.

16. Jacobs, R.M., Podolsky, L., and Okun, A.M., "Certification, standards and Quality Assurance", presented at the E.O.Q.C., Moscow, U.S.S.R., June 1971

17. Van Cronkhite, J.B., "Vision East; Soviet Standardization", Quality Progress, Vol. IV, No. 11, November 1971.

18. "Tabulation of Voluntary National Standards, Industry Standards, International Recommendations and Certification Programs for Consumer Products", NBS Technical Note 705, U.S. Department of Commerce, December, 1971.

COST-TO-PRODUCE

By
R. E. Biedenbender
Special Assistant, Cost/Value Engineering
Office of the Secretary of Defense (Installations & Logistics)

Introduction and Summary

In the past decade the Department of Defense has initiated numerous specialized programs aimed at effecting economy in defense. PERT-COST, Life Cycle Costing, Standardization, Value Engineering, Cost Reduction, Reliability, and Maintainability are examples. Such programs have, in many instances, achieved a modicum of success. Despite this, acquisition and support costs of defense systems have generally continued to rise.

In the light of this fact, and the fact that defense expenditures seem to be leveling off, the DoD has recognized that a fresh approach to cost is necessary. Such an approach was established in DoD Directive 5000.1, "Acquisition of Defense Systems". This Directive made cost of acquisition and ownership a principal design parameter.

This paper reviews three areas:

(1) General trends in DoD budgeting and system costs - and their implications.

(2) Concepts and activity in the area of cost-to-produce.

(3) Some implications of the above on specialized programs such as VE.

Budget Limitations

Most experts generally agree that defense expenditures are generally leveling off. The FY 73 budget is 6.4 percent of the Gross National Product (7.0 percent FY 72) - a 22-year low. In FY 68 defense was 39 percent of the budget - in FY 73 it is 30 percent. During the same period, the human resources share of the budget went from 32 percent in FY 68 to 45 percent in FY 73. Inflation has taken its toll. In constant dollars the defense budget is 30 percent below its FY 68 peak - and 8 percent below the FY 64 level.

Manpower and related costs are rising - from 42 percent in FY 68 to 57 percent in FY 73. From FY 68 to FY 73, military and Civil Service manpower has dropped by 1,440,000 - while pay and related costs have risen by almost $11 billion, or 33 percent.

Under a fairly constant defense budget, these trends have meant less money for research and development, and less for acquiring new systems to provide the standing capability necessary to fulfill U.S. defense commitments. This trend has been reversed to some degree during the past several fiscal years. From FY 64 to FY 71, for example, manpower and operating costs rose $7.6 billion. From FY 71 to FY 73 they decreased $5.8 billion. From FY 64 to FY 71, research and investment decreased $5.5 billion. FY 71 to FY 73, such spending will increase by $4.2 billion. Nevertheless, manpower and operating costs will continue to be a major portion of the defense budget.

Production and Support Costs

Now let's look at some specifics in the areas of acquisition and maintenance - the driving force in logistic support. The Department of Defense currently reports progress on some 45 major weapon systems to Congress. In March 1972, 31 of these had cost overruns. Some systems are now projected to cost over three times their original estimates to complete.

Even more important to our defense capability, in the long run, is the rise in typical unit development and production costs from one generation to the next. On 13 major systems, for example, development costs of the present generation has risen over 5 times, and production unit costs over 4 times. One senior official has stated, that at this rate, in just 40 years the entire Air Force budget will be spent on one plane; the Army budget on one tank; and the Navy budget on one ship. In contrast, performance growth as exemplified by factors such as payload, range, speed, avionics, and delivery accuracy, have risen on an average of only 1.8 to 3 times. Thus cost growth is rising more rapidly than performance. This rise is frequently diminishing the quantities DoD can buy. The quantity reductions on the C-5 and F-14 are examples.

The picture is similar in maintenance. Today the DoD services, repairs, overhauls, and modifies more than $100 billion worth of defense systems and equipments. Maintenance employs some 1,350,000 personnel, representing every field of technology. One-half million of these people are Department of Defense or contractor civilians. There are 20,000 maintenance shop facilities in 2,000 locations, totaling almost 300 million square feet, representing an investment of over $3 billion. Most important, maintenance costs are increasing, rising from $11.5 billion in FY 62 to a conservative estimate of $20 billion in FY 72, an increase of almost 75 percent.

Cost-to-Produce

Facts such as these have created top management recognition in the DoD that past attitudes on cost must change. As Dr. Foster told Congress early in 1972 "Classically the entire research and development community - and the military themselves - has favored performance over schedule, and schedule over costs. During the coming year, we plan to concentrate even more on readjusting these priorities until cost becomes as important as performance, and schedules are delayed to accommodate both."

He went on to say that during the coming year, areas to receive emphasis included the following:

(1) "Greater concentration on the use of production unit cost as a basic design parameter during concept formulation and engineering development."

(2) "Better incentives to encourage elimination of marginal requirements in system development that contribute more to cost than to effectiveness."

(3) "Greater stress on the achievement of high reliability and maintainability - and on demonstrating it in the test and

evaluation phase."

Many actions indicate that, particularly in terms of cost-to-produce, the Office of the Secretary of Defense means what it says. For example:

- A production ceiling price of \$1.4 million has been set on the close support AX aircraft.

- The Secretary of Defense has ordered that a possible replacement for the C-130 cannot exceed \$5 million - although many seasoned engineers estimate \$10 million a reasonable price.

- Because the costs of DD-963 destroyers is running about \$90 million a piece (up a factor of two), the Navy is asking for a new design called a patrol frigate, with a target price of \$45 million.

- The Army Main Battle Tank has been cancelled because Congress judged it still too expensive.

- The Army search for a replacement to the Huey helicopter has a unit price of \$600,000 established.

These are indications that unit cost-to-produce will be taken much more seriously by the design community in the future than in the past. Granted that cost-to-produce (or design-to-a-price) has arrived, what does it mean? Broadly speaking, the term cost-to-produce is used to denote control of the future production cost of an item during development.

Can cost-to-produce "work"? Yes it can. The challenge of cost-to-produce is not as new as it may seem. Let's compare the steps involved in the development of many commercial products with those of most military products. Figure 1 shows this process to be generally comparable with one important exception - control of cost-to-produce during design and development. Many commercial corporations employ cost-to-produce as a principal design parameter. Some defense contractors and internal DoD organizations have experimented with the technique. Cost-to-produce has been the subject of discussion in the VE literature since 1962. This work has languished primarily because of the low priority given it in the past.

COMPARISON
Figure 1

Commercial Product	Military Product
1. Conceive Idea	1. Express Requirements
Project sale price Research the basics	Budget cost-to-acquire Research the basics
2. Design and develop	2. Design and develop
Regulate cost-to-design Control future cost-to-produce	Control cost-to-design
3. Manufacturing	3. Contract for Manufacturing
Control quality Hold cost to standard	Inspect quality Audit actual costs
4. Deliver to Customer	4. Deliver to User
Instruct in use Maintain warranty	Train in use Provide logistic support

What principles should be followed to maximize the chance of a successful cost-to-produce effort? The following elements appear essential:

(1) Top management support and priority.

(2) Good initial estimates.

(3) Feedback on progress, to both the designer and to management (including the customer).

(4) Contract incentives, whenever possible.

(5) Inclusion of attainment of reliability and maintainability requirements as prerequisites of cost-to-produce goals.

Let's discuss each of these briefly in order. Top management priority and support is an obvious necessity. Without top management concern and interest, the best cost-to-produce system in the world would at best produce marginal results. The history of many specialized programs illustrates this fact.

Good, rational, initial estimates are necessary if requirements are to be meaningful, and if a host of other practices, such as "buy-in", are not to abort cost-to-produce effort.

Third, feedback on progress is essential. Neither management nor the designer, whose decisions largely determine future production and support costs, can take effective, timely action without cost visibility. An essential aspect of cost feedback is that to the customer. Frequently his requirements are marginally cost effective or dictate means whose uses are not essential to attainment of the military requirements. Development funds are scarce. Cost visibility of future production costs must be obtained while sufficient development funds are still available to take corrective action.

The use of contract incentives to motivate contractor design teams to meet cost-to-produce goals is an obvious corollary to the new priority of cost-to-produce. While incentives have been used for years, most of the time these applied to performance on the current contract - rather than expected cost performance in future contracts. Making part of the development contractor's fees subject to satisfaction of cost-to-produce requirements is a logical step consistent with current DoD objectives and priorities.

However, cost-to-produce cannot be made an end in itself. Certain technical or performance parameters must be met. Similarly, if DoD's objective is minimizing life cycle costs, attainment of minimum requirements for reliability and maintainability should be preconditions for eligibility for cost-to-produce incentives or fee awards. A recent study has shown, for example, that we miss our reliability goals in avionics by a range of 3 to 1 to 10 to 1. We must do better.

Let's turn now to some of the problems, alleged or real. First of all there are problems in definition. Although the general intent of cost-to-produce or design-to-a-price is understood, its meaning in practice must be quite clear if major errors are to be avoided. For example, does design-to-price mean that we will pay only a specific amount for an item, regardless of performance? Suppose two companies are competing in prototypes, and Company A meets the price requirement. Suppose Company B's price is one percent

higher, but its product demonstrates far superior performance and reliability. Would it be logical to award production to Company A? I doubt it. In the long run, cost effectiveness and total cost of ownership are far more important than initial price. It is important, therefore, that definitions be given precise meaning before contract negotiations begin.

We often hear that cost-to-produce "can't be used because there are too many unknowns." This objection is diminished by current policy for continuing assessment of technical risks, delay of full scale development until solutions are available for areas of major technical problems, and better test and evaluation prior to production. Greater use of parametric estimates, which can implicitly include consideration of such unknowns, should improve the accuracy of initial predictions. Lastly, initial cost-to-produce estimates need not be totally sacred. Opportunities exist for estimates to be appropriately modified as design progresses.

Other difficulties can be cited. How is escalation handled? To what level of design indenture should cost feedback occur? Answers to such questions exist or can be developed with time. What is important is not the ability to precisely predict future production costs before development begins, but the ability to track the evolution of cost-to-produce throughout development to permit initiation of appropriate action earlier than in the past.

Value Engineering

Cost-to-produce and cost-to-support as principal design parameters are a logical extension of current policy in DoD Directive 5000.1, Acquisition of Major Defense Systems. It unites the objectives of the engineering production and logistic support communities as never before. It provides a logical framework for many of the specialized programs - such as Should Cost, Standardization, Reliability, Maintainability, and Value Engineering - whose objectives in the past were inconsistent with overall program management objectives and priorities. This has frequently resulted in lip-service and worse abuses.

Let's take one of these specialties - Value Engineering - and examine its future in the light of current defense cost policy. If a layman on the street examined VE in the light of current cost concern, he would think that VE should be enjoying unprecedented popularity. We all know this is not always the case - while many of the past criticisms have been corrected - the "after-taste" lingers on in many quarters. However, the OSD is continuing to support Value Engineering. A number of actions have been completed to reform and revitalize the program. ASPR provisions are being simplified and improved.

Let's look briefly at some statistics on one aspect of the DoD VE Program - Value Engineering Change Proposals (VECPs) - to see why significant opportunities exist for defense contractors.

- Over one-half billion dollars have been saved by the DoD since the creation of Value Engineering Incentives.

- FY 71 estimated savings to DoD through VECPs reached $95 million - an all-time high.

- About 30 cents of each dollar saved goes to defense contractors.

- Several defense contractors report as much as one-third of their annual profit to be from VECPs. For example, five contractors averaged $3 million each in FY 69.

- This opportunity exists for smaller contractors as well.

As cost pressures increase, and these opportunities are made known, more government program managers will recognize that VECPs are consistent with their own objectives - they help reduce costs and prevent or reduce cost overruns. They frequently have secondary benefits, such as better performance or reliability. The Air Force F-15 and Maverick SPOS are excellent examples of progressive attitudes. Both have approved numerous VECPs estimated to be worth over $30 million. This illustrates that contractors who do their homework properly can achieve a mutually beneficial VECP relationship with DoD program managers.

Let's look now at some cost-to-produce and VE contractual interfaces. The first conclusion most people generally reach is that a strong cost-to-produce contract requirement during development is a better approach than a VE Program Requirement. This is so because cost-to-produce is a more direct approach to the basic problem. This does not necessarily eliminate VE. First of all the wise contractor with a strong VE program will use it to help meet cost-to-produce requirements. Secondly, an incentive clause can still be used. Experience on the F-15 and Maverick indicates that development costs can be reduced by this approach.

A strong cost-to-produce program should enhance benefits from VE contract incentives on future contracts. Cost feedback will better identify specific areas of future VE opportunity. A strong cost-to-produce requirement should also alleviate fears of some DoD personnel of contractor design malmotivation - deliberately costly design to achieve future VECP savings. In the long-run, cost-to-produce should make VE a more viable, meaningful program for both the practitioner and management.

Summary

In summary, overall defense economics currently dictate much closer control of future production and support costs during the design and development process. As a result cost-to-produce, which has been a major design parameter commercially, is now being viewed similarly in defense work. Timely feedback to the designer and the customer is a key element to successful cost-to-produce work. Cost-to-produce results can also be a significant factor in development contract fees.

The top-down cost-to-produce and cost-to-support approach provides improved motivation for proper employment of specialized programs, such as VE and reliability, which are economically oriented. Future emphasis on cost-to-produce and cost-to-support should contribute to more effective use of such specialized programs.

SOME ECONOMIC ASPECTS OF AVIATION SAFETY

HANS W. WYNHOLDS
Lockheed Missiles & Space Company, Inc.
Sunnyvale, California

This paper represents an initial attempt at addressing the fundamental issues of safety and regulation by constructing a simplified model of an industry that produces a single service (transportation) for which the cost of production and the value of the service are precisely known. The issue of just which level of safety promotes the public interest best can then be addressed in a case in which the quantitative results of taking alternative regulatory schemes are available.

It is shown that a fundamental logical problem requiring societal value judgments must be solved in order to determine the necessity for regulation and, if any is imposed, the type of regulation that is socially optimal. In other words, there are no simple answers even in a simplified case. This result suggests that considerable care will be required in order to arrive at an optimal social policy in more complex problems, such as those presented by the interaction of the regulated airlines and the relatively unregulated aircraft manufacturers. It also suggests that highly simplified models may be useful in delineating some of the issues in these more complex and realistic problems of safety and regulation.

Introduction

One of the most important issues in commercial aviation is the assurance of optimum safety for the public, passengers and personnel; and the capability of the industry to respond to this need. An economic system dealing with aviation safety consists of three basic elements: the airlines, the aircraft manufacturers, and the public—principally represented by the airline passengers. This economy is operated through the exchange of money for products and services. The passenger pays the airline for the trip, the airline pays the manufacturer for the aircraft, and on rare occasions the manufacturer pays the passenger (or his estate) for damages caused by an unsafe vehicle (either by design, manufacture, or operation).

Inherent in the purchase of any of these goods is the assumption of some level of risk, or inversely, some level of safety. Intuitively, the cost of that safety is included in its total price. It is then argued that passengers can effectively value their desire for safety as an integral portion of the service received and, in an economic market sense, determine that level by selecting alternative modes of transportation over that route, or by adjusting the demand by foregoing the trip altogether. Even damage payments by the manufacturer may be considered as the reimbursement of purchase price, if you will, to the public for the safety that was not supplied with the original product or service. Thus, if the economic model could be constructed and evaluated, then perhaps an insight could be gleaned of an industry-wide optimum safety level. This could, in turn, translate directly into product specifications and operational safety standards.

The central economic problems of any society are concerned with what to produce, who should produce it, and who should consume it. Such judgments can be carried out with differing amounts of government control, depending on the economic philosophy of the society. Therefore, to determine whether it is necessary to regulate an industry at all, and if so, how it should be regulated, requires that the nature of the necessary choices be understood.

A fundamental question that is suggested by this problem is whether or not the basic process and fact of the regulation impedes the capability of the industry to meet changing demands rapidly and effectively. If so, what are the basic objectives and purposes of regulation and what are the alternatives? Could the commercial aviation industry be regulated in a way that might be less cumbersome than the present system that might meet the basic objective of promoting the public interest equally well?

The Economic System in Review

"Safety, to a large extent, is a purchaseable commodity. This is not to imply that improvement is simply a matter of spending more money. The intention is to emphasize the very large degree to which air safety is subject to control, and that a primary control is economic."[14] To mimic Schelling[18], safety is indeed different from most consumer goods, and its purchase different from most commodities.

A review of the market structure of the aviation industry is similar to the market structure of aviation safety (Figure 1). The passenger pays the

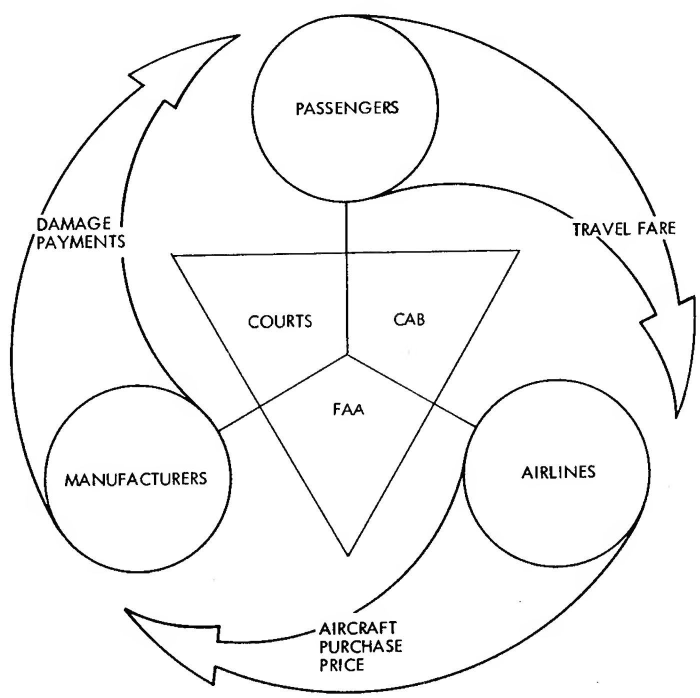


FIG. 1 MARKET STRUCTURE FOR AVIATION SAFETY

airline a fare for the privilege of being transported from one place to another. In return, he assumes some degree of risk. The level of risk is partially influenced by the way the airline operates its vehicles and partially by how well the manufacturer designs and builds his vehicles. The extent to which the passenger subjects himself to this risk is addressed in many of the reports on aviation safety statistics.

In turn, the airline uses a part of the passenger's fare to purchase a fleet of aircraft from the manufacturer. One risk assumed by the airline is the possible loss of revenue due to inadequate vehicle design. If the craft is unreliable, the nuisance factor may be enough to cause concern for the airline. If the planes operate unprofitably, the airline may wish it had never purchased them and look to replace them as soon as is feasible. But if an airplane is unsafe, and has demonstrated this fact in operation, the airline will lose substantial revenues due to:

1. having lost the value of that vehicle (and associated personnel),
2. not having that vehicle in service to continue earning for the company,
3. loss of passenger confidence reflected in selection of alternate service.

The manufacturer, in some instances, will make payments to selected passengers (or their respective estates) for damages incurred through malfunction of its product. Thus, the manufacturer assumes the risk of making and selling a defective item and having it cause injury or death. If the product were simply undesirable, it might not sell. This could cause serious problems with manufacturers who invest a substantial portion of their resources into a relatively few products, such as large aircraft.

Undesirability of its products might slowly drain a compay of its resources. But large scale legal action and lack of follow-on orders because of unsafe design could turn that trick overnight.

It is clear that the ultimate risk for each of the parties is rather extreme. It should be noted, though, that the financial transactions are not at all similar for any of them. On a rather continuous basis the passenger pays relatively small amounts of money to the airline. Periodically, the airline invests in new equipment, but it must pay a handsome sum when it does. And occasionally the manufacturer must settle an astronomical legal claim. By financing the aircraft and by insuring its safe operation, the airline and manufacturer attempt to spread out these large payments into smaller ones over a longer period of time.

Each of the transactions between participants in the aviation market is influenced by regulation or control. As the illustration indicates, the Civil Aeronautics Board (CAB) regulates the pricing of fares and competition on routes of airline operation. The Federal Aviation Agency (FAA) influences the design standards of the manufacturers and the physical operations of the airlines. Finally, the judicial system with its "tort" law, by way of its legal structure and operation, controls the penalty payments that are paid by the manufacturers to the injured passengers.

Regulation

Reviewing with the perspective of economics, in 1938 the Civil Aeronautics Act pulled together, for the overall purpose of economic regulation, legislation which had previously controlled air carriers. Air carriers then in existence were granted "grandfather" certificates for routes then being operated. However, the 1938 act contained provisions for the enfranchisement of new entrants into the industry, and for the certification of new and additional routes. The basic criteria for new routes were a finding of public convenience and necessity, and a determination that the applicant for a route be deemed fit, willing and able to service the route.

When the CAB was established 34 years ago, it fulfilled several urgent needs of the time. The struggling airline industry needed the security of federal operating certificates; the public needed protection from marginal, unsafe operators; the postal service needed dependable schedules and assured capacity over designated routes; national security considerations required a healthy and growing air transport civil reserve fleet. It was President Roosevelt who called for legislation establishing a federal regulatory agency modelled after the Interstate Commerce Commission.[3]

The fundamental economic philosophy underlying the Federal Aviation Act of 1958 is little different from that employed in regulation of other public utilities, or public services having a public utility flavor. Regulation is supposed to balance the wastes of unbridled competion against the evils of monopoly.

Air transportation in this country has followed the concept of regulated competition. Theoretically, this produces enough competition to guarantee the public a choice among strong and stable air carriers while, at the same time, protecting these carriers against hit-and-run, unregulated competition.

Under the Federal Aviation Act, the CAB is empowered to prescribe minimum and maximum rates for domestic air carriers. The intense competition in the industry makes continued control over maximum and minimum rates less desirable and less necessary than they have been in the past. However, in the absence of free entry, and in view of the minimal degree of price competition, competition alone is not sufficient to protect the public against excessive rates. And, says Billyou, "continued control over minimum rates is also required to prevent unduly low rates which, in usual competitive circumstances, might result in service deterioration and a lowering of safety standards."[2]

Part 601 B of the Federal Aviation Act of 1958, in calling for the highest degree of safety in airline operation, fails to define whether industry or the FAA is responsible for achieving this. The operating environment is, of course, provided by government. Industry is responsible for producing aircraft, techniques and management to operate safely in this environment. And the ultimate responsibility for safety is borne by the individual working as part of the system.

Risk

To determine more specifically where responsibility lies, it is first necessary to make the distinction between public safety and private risk. Dr. Warner, former vice chairman of the CAB, stated his philosophy in 1936: "There should be a limit on government responsibility, and I suggest the reasonable limit is that the government protect anybody who is too helpless to protect himself." This line of reasoning would seem to place the prime responsibility for safety with the government for public carriers and with the individual for private aviation.

This philosophy apparently stands intact today. Starr[19] categorizes societal activities as those in which the individual participates on a voluntary basis, and those in which he participates involuntarily. In

the case of "voluntary" activities, the individual uses his own value system to evaluate his experiences and this is likely to represent, for that individual, a crude optimization appropriate to him. "Involuntary" activities differ in that the criteria and options are determined not by the individuals affected but by a controlling body. Such control may be in the hands of a government agency, a political entity, a leadership group, an assembly of authorities or "opinion makers" or a combination of such bodies.

There exists a separation by several orders of magnitude between the voluntary and involuntary societal acceptance of risk. As one would expect, we are loath to let others do unto us what we happily do to ourselves. Starr also points out that the disease rate appears to be an individual's subconsious yardstick against which he measures the acceptability of risk on a voluntary basis.

The risk position of commercial aviation is partly voluntary and partly essential, and additionally is subject to government administration as a transportation utility. It is now approaching a risk level comparable to that set by disease, but increased public participation will undoubtedly increase the pressure to reduce this risk.

Responsibility

That you should pay for damage you do to someone through your fault is the basic principle of our "tort" law. The fact and extent of liability are determined by the law of negligence. Liability involves injury to passengers, injuries to the public, and damage to property. As such, a manufacturer of aircraft products is liable for harm to others caused by a design defect in his products.

Virtually any design decision or judgment it makes is subject to challenge, and the challenge might come in an adversary proceeding where the ultimate judgment is made not by engineers, but rather by fact-finders who have no technical background.

However, the degree of legal care required of a common carrier may be greater than that for a private individual. Legal responsibility is often associated with express or implied warranties to the buyer that the product has been made in accordance with the purchase agreement, type specifications and regulations. As matters stand, however, FAA certification not only fails to insulate a manufacturer from liability of alleged design defects, but it may even expose the government to liability of its own.

Today, the liability of a manufacturer extends far beyond liability based on his negligence. Under the doctrine of strict liability, an injured party need only show that the product was defective when it left the manufacturer's hands, that the defect later caused him injury or harm, and that in the interim there was no substantial change in the product. Negligence need not be shown, for liability attaches even though the manufacturer has excercised all due care.

Almost without exception, a manufacturer will be held pecuniarily liable for damages caused by design defects in his products. Many manufacturers have first become aware of this when judgments have been rendered against them in injury and death cases. The lesson has sometimes proved very costly.

The damages sustained by the plaintiff depend, generally, on the degree of the injuries suffered or, in case of death, monetary loss. Damages for death depend on the standards accepted by the particular law involved. In the U. S., each state has its own "wrongful death" law. Most do not have any artificial limitations of damages, and fix damages according to the "pecuniary loss", or money loss, sustained by the survivors of the decedent.

Of course, it is for the defendent to show the uncertainties and frailties of life, and thereby minimize damages. The contested question of potential will be resolved by the jury at the trial. In practice, many cases are settled before trial.

Review

It has always been, according to Schelling[18], that "the avoidance of a particular death—the death of a named individual—cannot be treated straight forwardly as a consumer choice. It involves anxiety and sentiment, guilt and awe, responsibility and religion. Yet, when we ride an airplane, death is about the only risk that we consider."

He contends that in an already advanced economy, many of the ways of reducing the risk of death are necessarily public programs, budgetary or regulatory. In fact, he adds, safety regulations must be partly oriented toward guilt and responsibility. And if it turns out that safety is a public good and not everybody wants it at the price, or that the tax system will not distribute the costs where the benefits fall, so that we are collectively deciding on a program in which some of us have a strong interest, some a weak interest, and some a negative interest, that makes it rather like any budgetary decision that the government makes.

If, on the other hand, it is assumed that safety is a commodity which can be reliably priced, then an argument could be made for reducing the present stranglehold regulation on this system. Even Schelling concedes that moral judgments are fine; but that in the end it may be the passengers who want more safety and who should bear the cost.

In review, it is significant to note that this economic structure of aviation (and safety) is a "closed loop." This implies that there exists an economic feedback channel within the system. However, with external regulation and control, these forces attempt to reduce the dynamics of the system and to gradually introduce new forces without creating major disturbances. As is so often the case, unfortunately, controllers tend to extend their authority (or at least not to lose it) and they tend to be slow in reacting to new forces.[2,6]

The Elements of Decision

Government provision of facilities to accelerate the development of air transportation specifically, and aviation generally, is a responsibility imposed by the Federal Aviation Act's promotional requirements. Subsidy, direct or indirect, is a recognition by Government of a responsibility to provide facilities for a common use, so long as the public interest is served.

The advantage to a beneficiary of Government promotion and the use of public funds includes the possible disadvantage of continuing Government control. When an air carrier is on subsidy, it must expect that the managerial discretion permitted will be more limited than for unsubsidized carriers. Air services on a subsidized basis should be authorized, but only where such services promise to produce economic, social, or national security values that clearly exceed the cost of subsidization.

To determine the conditions under which regulation may be desirable, a simplified structure of the aviation economy is proposed. In it a single service is provided by an industry which also produces the goods

necessary to supply it. In other words, the original feedback loop structure would reduce to a dual entity system with the merging of the manufacturer and the airline. It is also assumed that resource allocation in the society is ruled only by concern for the public interest.

Most early attempts at determining an optimum safety strategy focused on the costs of safety and of accidents. The model proposed by Tye[20] and Canale[5] in Figure 2 suggests that the total "safety" cost (TC)

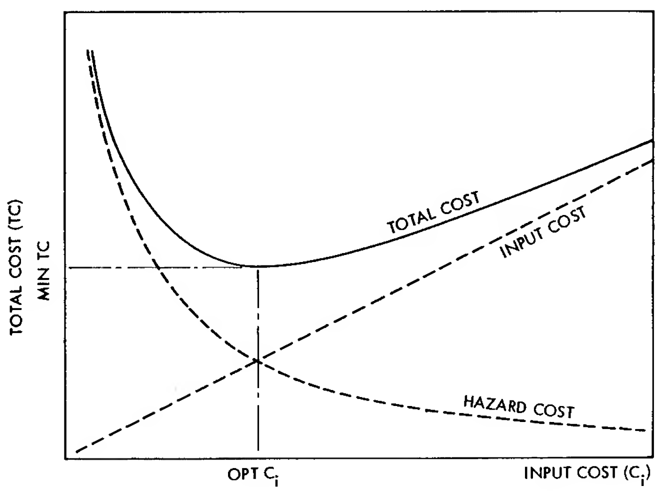


FIG. 2 COSTS DUE TO SAFETY AND HAZARD EXPOSURE

is comprised of the cost of building in safety (input cost) and the cost of not having built in safety (hazard cost). By not having been designed, built or operated safely, a product could expose people to its hazards. The less the input safety, most likely, the greater the hazard level. Thus, hazard cost may be assumed to be a function of, and inversely related to, input cost. The hazard cost might even be assumed to be monotonically decreasing with respect to input cost and, for large input costs, asymptotic to some level which can be attributed only to unavoidable risk. The resultant sum of costs would then be an everywhere convex function implying that a global minimum total safety cost exists for some optimum input cost value.

One informed opinion is that the present level of safety input is far less than whatever the optimum might be. The argument is that the entire development program of an aircraft may not have cost more than the ultimate cost of a single major accident. For example, Kubli[12] estimates the total exposure of a Boeing 747 at over $84 million for one accident other than a midair collision, and at $35.5 million for a DC-8.

A principal difficulty with determining the amount of resources actually allocated to safety is that it must be ferreted out from normal design costs, reliability cost, and the like. How often is a reliable design also a safe one, or an efficient operation also one which reduces exposure? Estimates of the input cost of safety for some recent programs have been anywhere from 0.5 to 2.0 percent of the total engineering budget.

The fundamental weakness of this model is that it does not perceive the system as a closed loop. Rather, it assumes certain costs of accidents as contributing to the total cost of safety; whereas these costs will actually fluctuate in the courts (the proverbial marketplace of product safety) as an expression of societal or consumer preferences.

A purely economic analysis would consider not only the desired level of safety, but also the market form in which it is to be achieved. If perfect competition would not produce the desired level, and some form of monopoly could, the question of regulation then becomes apparent.

A first step toward solution is to determine the cost to society for varying levels of safety. These costs should be measured in terms of the net demand placed on the factors of production in the economy and should include the costs due to externalities. In principle, the cost of regulation should be included but, in practice, these costs are so small as to be negligible.

The social revenue function indicates how much society is willing to spend for a given level of safety and is a function of the demand curve. Put in other words, this is the amount of resources that would be diverted from other types of consumption or production in order to obtain the desired safety level. Within the simplified structure of this model, these preferences are continually expressed in the form of fares and legal action.

By knowing the resources required to attain a given level of safety and the resources that this can command (reserve), the social profit can be constructed for any level as the difference between them (Figure 3). Any good that has a positive social profit

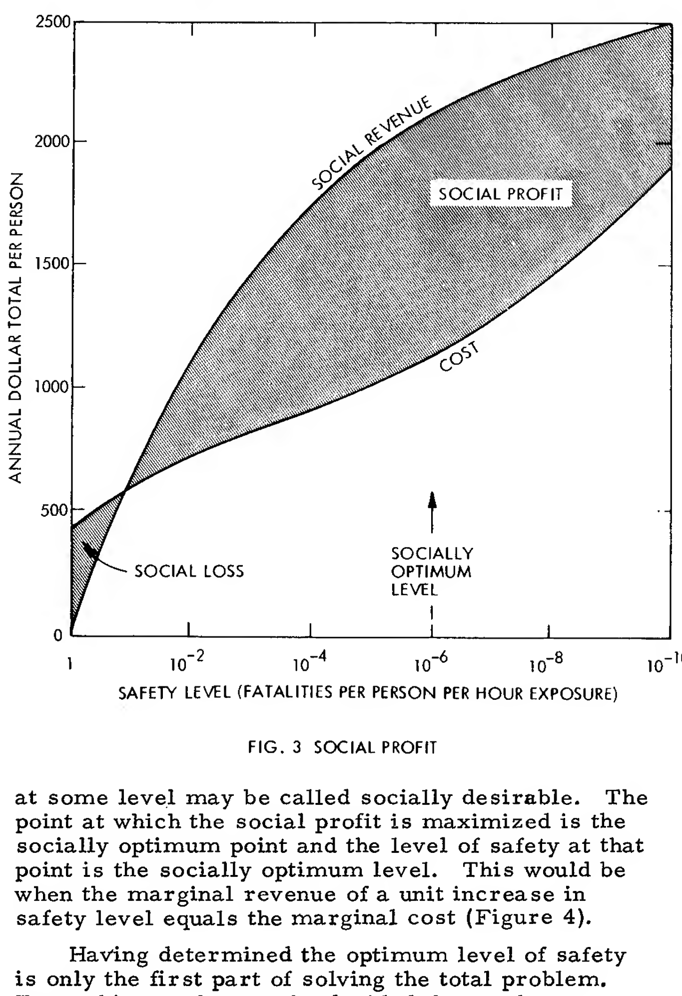


FIG. 3 SOCIAL PROFIT

at some level may be called socially desirable. The point at which the social profit is maximized is the socially optimum point and the level of safety at that point is the socially optimum level. This would be when the marginal revenue of a unit increase in safety level equals the marginal cost (Figure 4).

Having determined the optimum level of safety is only the first part of solving the total problem. From this result must be decided the market environment which will best assure the desired outcome. Several alternatives are reviewed in the manner of Howard and Matheson[10].

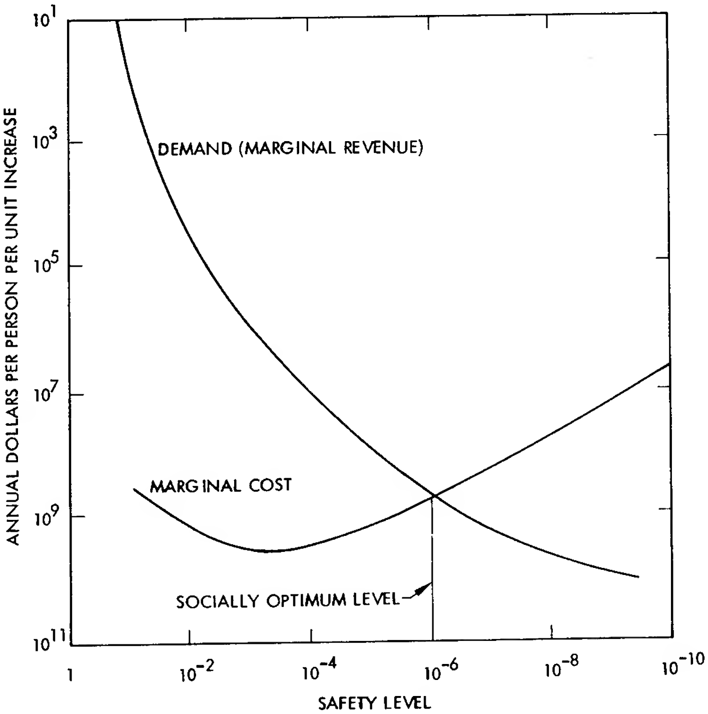


FIG. 4 DEMAND (MARGINAL REVENUE) AND MARGINAL COST

One market form would be to grant a monopoly and allow discriminating pricing. Translated, this means that each person would be charged a price for safety equal to his value of it. If this strategy were followed, the revenue to the monopolist would be the same as the social revenue and his profit just the social profit.

A characteristic of this arrangement is that the price of the last incremental increase for safer design as determined from the demand curve must be equal to the marginal cost of producing that safety. This provides proper economic signals but requires that the monopolist have considerable wisdom in determining just what price each individual is willing to pay. In practice, this is usually not feasible or even desirable. But on occasion, "skimming" is used as a pricing strategy.

Another feature of the monopolistic arrangement is that the social profit accrues to the holder of the monopoly, rather than to society. The undesirability of this result is a function of the general economic environment and the nature of the monopolistic enterprise, be it public or private.

If a non-profit public corporation were allowed to do discriminating pricing, the revenue and costs would be just the ones discussed above and the organization would be hard pressed not to be profitable. If the average revenue and cost per incremental increase in safety are computed, then the case where average (and consequently the total) profit are zero, is usually in excess of the socially optimum level. Therefore, a non-profit monopoly allowed to do discriminating pricing does not necessarily achieve the social optimum.

One modification could be to allow partial price discrimination, charging a high price for safety to some users and charging the rest at the marginal cost at the socially optimum level in such a way that total profit is zero. The result would be a self-sufficient enterprise that operated at the socially optimum safety level. However, the problem of price discrimination remains: who should pay which price?

In the interest of fairness, the alternative of general pricing is considered. In this case, the revenue of the enterprise, now called the general revenue, is just the number of users (or passengers) times the price indicated by the demand curve. Since there are individuals who were willing to pay more, the general revenue will be less than the revenue produced by discriminating pricing (Figure 5).

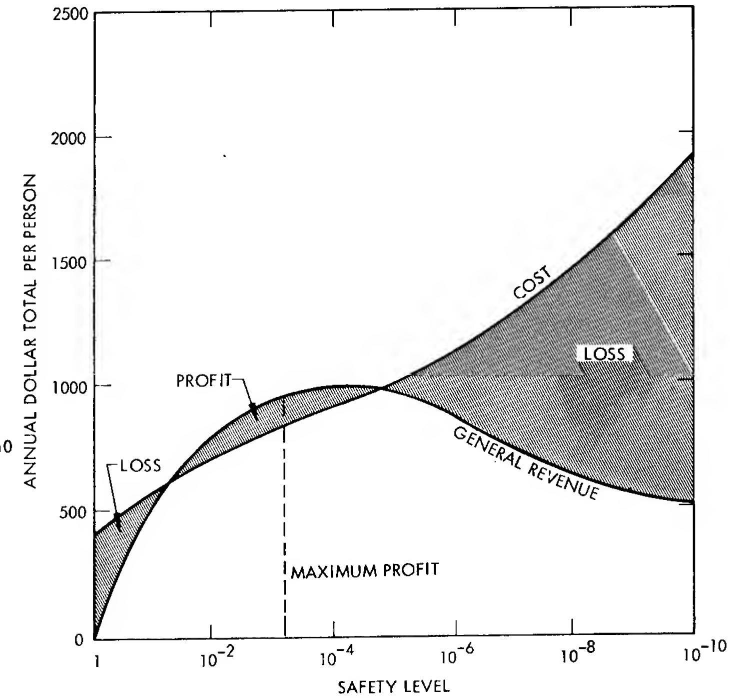


FIG. 5 GENERAL PRICING

Thus, if the enterprise is allowed to maximize profit given only that it must establish a general price, it will most likely produce a level of safety lower than is socially optimal (Figure 6). The basic

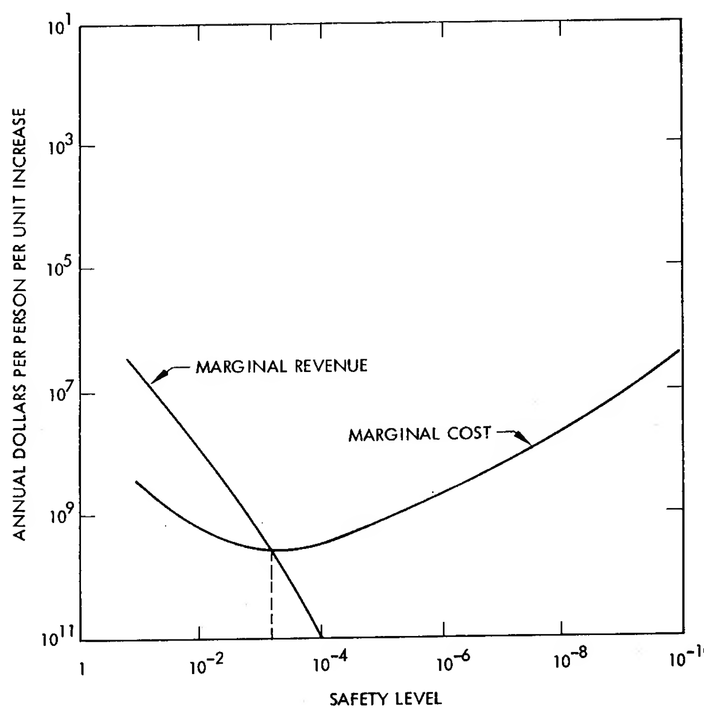


FIG. 6 MARGINAL REVENUE AND MARGINAL COST UNDER GENERAL PRICING

difficulty is that the marginal revenue is far above the marginal cost. Therefore, individuals who are willing to pay more than it would cost to improve the level of safety but less than its price will not be served even though it would increase social profit to do so.

If the monopolist were required to set the general price equal to the marginal cost at the socially optimum level, the enterprise would probably operate at a loss because of a shift in the marginal revenue function. The maximum social profit would be achieved, however, and retained entirely by the customer (one extreme of wealth distribution). The loss for this kind of arrangement could be paid, perhaps out of the general taxes of society. But it would be difficult to devise a fair taxation scheme, and one that would leave unchanged both the social revenue and cost functions, hence the same social optimum. It appears that payment of the subsidy through taxation is not a solution to the problem, rather a transformation of it into another possibly more difficult form.

A non-profit enterprise forced to do general pricing would require that the level of safety be determined by operating with average revenue equal to average cost (Figure 7). While this level does insure

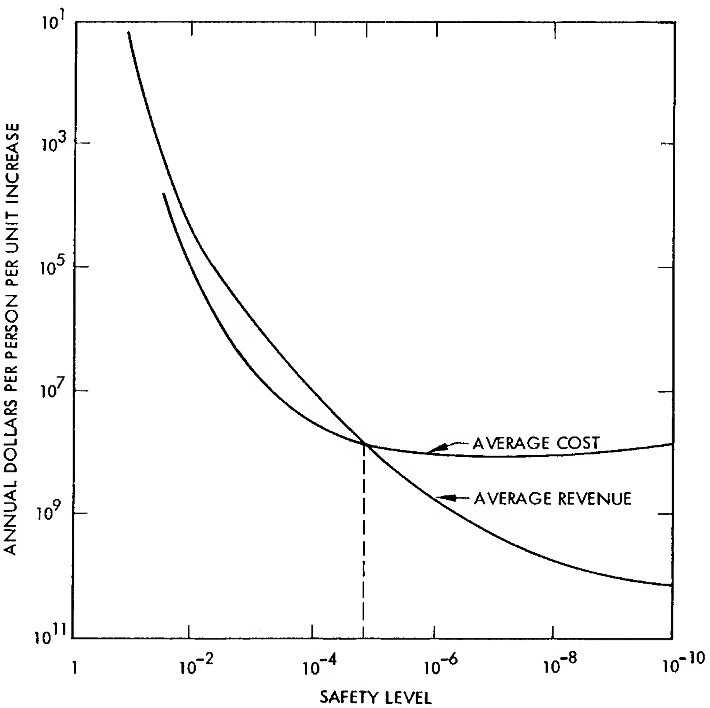


FIG. 7 AVERAGE REVENUE AND AVERAGE COST UNDER GENERAL PRICING

self-sufficiency, it most likely does not maximize social profit and will probably price safety at higher than its marginal cost.

An added dimension to this problem is the selection of capital investment level. The question of which among several technologies to employ will influence cost of improving safety. If there is some level of safety at which a particular technology produces minimum cost, it is an efficient technology. Therefore, instead of dealing simply with a cost curve as before, alternative technologies require the construct of a minimum cost curve.

Regulatory possibilities to assure the socially optimum technology in the public interest are similar to those proposed when only one was considered.

The monopoly with discriminating pricing has all the characteristics described in the previous discussion. Another scheme, general pricing at marginal cost with profit incentive, suggests establishing a general price equal to marginal cost for the selected technology and an incentive to maximize profit from general revenue. The tendency would be, however, to select a technology with a smaller investment than is socially optimum and producing a lower level of safety, thus driving its marginal cost (and price) higher.

The common method of regulation is general pricing with an allowable return on investment. By establishing a maximum profit on an investment, the tendency will be to select a high investment technology and then produce a relatively low level of safety. The company can achieve increasing profits by increasing investment as long as there is a level of production which generates enough revenue to cover cost plus the allowable profits. In fact, because profits are proportional to investment, there is an incentive to be less efficient and, if possible, to use an inefficient technology to produce even higher profits.

With a general pricing monopoly, the company is required only to sell at a general price and is then allowed to maximize profit. In this case, there is a tendency to pick a low investment technology and to produce a low level of safety.

Conclusions

The challenge is to provide incentives other than disaster, and before disaster, for insuring public safety. Projections by Lundberg[17] and Laughlin[13] estimate that by the year 2000 there will be between 9,000 and 18,000 passenger fatalities per year from transport aviation. This assumes the present rate of fatalities applied to the expected increase in activities.

Is there a tolerable accident rate? Some assert that the only acceptable accident rate is perfect safety[9]. This is an ideal objective. However, if this were completely adopted, it would make safety tradeoffs impossible. Compromises can be made only if one accepts some probability of trouble due to such compromise. A tolerable accident rate could be justified morally on the basis that the public knew it was accepting some risk in return for the advantages of air travel. This rate could be balanced against the loss of life that would occur if there were no air transportation, whereas a higher rate would exceed the public's confidence. The acceptance of a "tolerable" death rate may be morally repugnant, says Lederer; but what is the alternative?

Having constructed a simplified model of the aviation industry it has been shown that the characteristic which makes a product worthwhile for society is that there is a positive social profit for some level of safety. Being socially desirable, the socially optimum level of safety is defined as that level which maximizes social profit. This principle of maximization extends to the selection of technology, or investment level.

Several methods of delegating the investment and production level decisions to a lower authority were examined. All regulatory schemes considered are compromises among conflicting criteria. Some tend to encourage all socially desirable transactions, others to create self-sufficient enterprises or to distribute wealth "fairly," even to give freedom to the various economic agents. Even in the simplest of cases, however, the course of regulatory wisdom is not clear.

"The most effective competition is not always a question of quantity. Often it can depend on the strength of possibly fewer corporate entities competing against one another. A most important purpose of competition is to serve the public interest"[2].

Recognizing, however, that the system may not be operating optimally at present suggests the alternative argument for relaxing artificial control and allowing the system to regulate itself. This would not be the first time that a regulated industry, when turned loose, would operate more efficiently[21]. Professor Cherington, in a recent editorial, also suggested that the keystone of a new regulatory strategy be the "rapid opening up and spread of new services and markets"[6].

The May 1968 issue of Space/Aeronautics devoted its entire issue to "Air Safety." Excerpts from the opening editorial admirably summarize the situation:

"As a concept, air safety has no enemies. As a reality for the 1970's, though, it now appears a dubious prospect."

"It's clear that air safety is going to raise the fares. Higher fares mean fewer passengers—a prospect that carriers view with alarm at a time when they have gone deeply into hock for new, high-productivity equipment. Fewer passengers also mean fewer planes—a prospect that manufacturers similarly view with alarm after having extensively committed to building the new, high-productivity equipment."

"In the end, the proper level of safety can perhaps be decided only in the marketplace."[4]

Bibliography

1. Bear, M. G., "Safety Standards and Profitable Operations," presented to the American University, 7 November 1961.
2. Billyou, DeForest, Air Law, Ad Press Ltd., New York, N. Y., 1964.
3. Burkhardt, Robert, "The CAB—Should It Be Changed? Airline Management, March 1972.
4. Campbell, John B., "Safety for the 70's," Editor, Space/Aeronautics, May 1968.
5. Canale, S., "Hazard Risk Measurement and Optimization," presented at Government-Industry System Safety Conference, NASA, Washington, D. C., May 1968.
6. Cherington, Paul W., "The Future of Airline Regulation," Aviation Week & Space Technology, 28 August 1972.
7. Coie, J. Paul, "System Safety - Legal Aspects," presented at System Safety Symposium, Seattle, Washington, June 1965.
8. Fitzsimons, James M., "The Aircraft Manufacturer's Product Liability," Astronautics & Aeronautics, November 1968.
9. Hammer, Willie, "Numerical Evaluation of Accident Potentials," Annals of Assurance Sciences - Fifth Reliability and Maintainability Conference, American Society of Mechanical Engineers, New York, July 1966.
10. Howard, R. A. and J. E. Matheson, "A Preface to a Theory of Regulation," Report No. 7370B-6, Stanford Research Institute, Menlo Park, California, February 1969.
11. Kreindler, Lee S., "How Does Accident Litigation Affect Aviation?", Astronautics & Aeronautics, November 1968.
12. Kubli, A., "Air Transport Insurance and Reinsurance on the Advant of the Jumbo Jet and SST Era," Reinsurance, May/June/August 1969.
13. Laughlin, T. F., "Inherent Safety Design in Current and Future Aircraft," presented at the Orient Airlines Research Bureau Safety Seminar, Tokyo, Japan, November 1969.
14. Lederer, Jerome, "Air Safety - A Study of Ethics, Economics and Attitudes," Flight Forum, Winter-Spring 1964.
15. Lederer, J., C. O. Miller and C. F. Schmidt, "Economics of Safety in Civil Aviation (Planning Study)," Technical Report ADS-7, Federal Aviation Agency, Washington, D. C., December 1963.
16. Lowell, Vernon W., Airline Safety is a Myth, Bartholomew House, 1967.
17. Lundberg, B. K. O., "The Allotment-of-Probability-Shares (APS) Method, A Guidance for Flight Safety Measures," Aeronautical Research Institute of Sweden, presented at the International Symposium on Civil Aviation Safety, Stockholm, Sweden, April 1966.
18. Schelling, T. C., "The Life You Save May Be Your Own," Problems in Public Expenditure Analysis, The Brookings Institute, Washington, D. C., September 1966.
19. Starr, Chauncey, "An Overview of Public Safety Problems," ASSE Journal, July 1969.
20. Tye, Walter, "Unresolved Civil Airworthiness Problems," Institute of the Aerospace Sciences, Inc., 1959.
21. "Another Rail Way," Forbes, 1 March 1972.
22. "They No Longer Say, 'I Got A Lemon'—Nowadays, They Say, 'I'll Sue'", Forbes, 15 April 1972.
23. "Transportation-Untangling the Nation's Lifeline" Business Week, 26 November 1966.

TOTAL LIFE COST AND THE "ILITIES"

Stanley Duhan and John C. Catlin, Sr.
Tennessee Valley Authority
Knoxville, Tennessee

Many approaches have been taken in recent years to predict the total lifetime cost for any given system or installation. The military requires guarantees on acquisition cost, maintainability, reliability, and performance while industry is interested primarily in acquisition cost and performance and has an increasing interest in the cost of ownership of their facilities and the cost of warranties of their products.

Most are familiar with the "Assurance Sciences" as applied to military equipment. It is not so well known, however, that we in the Power/Utility Industry are also vitally concerned with the application of the "Assurance Sciences" - the so-called "Ilities" - to our installations.

In our industry we are vitally concerned at the onset of a new project with: the design and fabrication of equipment for the system, performance, support characteristics, and delivery lead time. All of these elements can be translated into dollars, and then, synthesized into a total life cost for the entire project. This paper presents a method of developing total life cost for: new projects, revisions to existing systems of facilities, or items of equipment or modules. The mechanics of this methodology may be exercised by man or machine. The cost effectiveness of machine manipulation of data is dependent on company capability and project complexity. The methodology must be adjusted for the industry involved and the type of item being analyzed. Total life costs are developed as follows:

1. Operational and maintenance studies are conducted.
2. Cost worksheets are used to document each study and to assure compatibility of studies conducted.
3. A system summarization of all the studies is prepared. This summary includes such items as operating personnel, fuel and utilities, and facility and construction costs.
4. Finally a project summary of all the system summaries is prepared to which project costs such as land acquisition, licensing, project management, etc., are added.

Use of this method requires input data from two sources--engineering design and the using organization. Engineering design must state the parameters to be expected from its design. The using organization must provide the expected cost structure for labor, supply, etc. Input data is divided into the following categories:

Identification of equipment
Reliability information
Repair data
Spares data
Supply
Maintenance labor
Support equipment

This input list must be tailored to the user's requirements. For instance, if doing a study for the military, such items as publications, pipeline spares, training, and transportation must be added.

The above input categories are divided as to source as shown in figure 1.

The use of this method starts with establishing the maintenance support concept for each of the various designs being considered to accomplish a specific task. (It is assumed that if a design is being considered, the manufacturer has agreed to meet all required operating parameters). For example, some designs may contain nonrepairable components while others may have repairable subassemblies. Once these maintenance concepts have been established, each design configuration under consideration is evaluated on the worksheets provided for this task (see figure 2). They are:

Cost Worksheet - Material
Cost Worksheet - Labor and Support
Cost Summary

The addendum to this paper provides a description of the columns on the cost worksheets and briefly explains the functional elements which are contained within each column entry.

When each design study is completed and the optimum design concept is selected, its support characteristics are entered on a system cost summary worksheet. This worksheet also contains other system characteristics such as facility costs, operating personnel costs, and fuel and utility costs. Figure 3 is a typical system cost summary worksheet. When completed, this worksheet will provide the predicted total life cost of a system. In addition to total cost, the worksheet will also provide:

System failure rate
Total maintenance man-hours
Man-hours/operating hour

As an added feature, the reliability input can be evaluated for safety effects. By studying the effects of the various failure modes, not only can the safety of the system be analyzed but also the economic effects of predictable minor accidents can be included. The output can predict the total costs of losses due to an accident such as downtime, personnel lost time, secondary effects, and insurance and warranty as well as the usual maintenance costs associated with the failure of a given part or assembly.

The final project summary may indicate that variations in maintenance time requirements have very little effect on total cost and on the final selection of a design configuration. The same applies to the other "Ilities." Therefore, it is cost over the total life of the design that must make this selection decision.

Figures 4 and 5 are presented to summarize the results obtained from the cost summary sheets of two representative studies which were made using the described methodology. The first study compared three different design concepts which could satisfy a design requirement. The second study compared three different maintenance concepts for a specific design requirement. In both cases it can be seen that maintenance is a small segment of the total life cost.

It is known that improvements in maintainability, reliability, system safety, and human factors engineering all tend to increase acquisition costs while reducing support costs. Therefore, it should be mandatory that the design engineer evaluate the relative importance of support costs to acquisition costs and then expend his engineering budget in the manner in which this evaluation dictates. This evaluation will vary widely with different types of equipment. It will also vary as the quantities on order increase.

In conclusion, the method described will provide the designer with the necessary information for selecting a design based on the one commodity the user is most interested in--dollars. If the user can get more operating hours per day from equipment, I am sure he will gladly expend one more maintenance hour per operating hour than planned for on that equipment because this means a cost savings for him. This does not imply that maintainability studies should not be conducted, only that a parameter other than man-hours per operating hour should be used to determine the value of each study. These same statements can be applied to the other "Assurance Sciences." In the final analysis the design that can operate within specification requirements and has the lowest predicted total life cost should be the design that is chosen.

Addendum

Cost Worksheet - Material

1. Part number - Manufacturer's part number.

2. Part name - Manufacturer and user accepted part name.

3. Quantity per system - Self-explanatory.

4. Failure rate per 1000 - This rate is the component rate and includes primary and nonprimary malfunctions.

5. Total failures/1000 operational hours (OH) - This number represents total system failures and is the product of the component failure rate (column 4) x the quantity per system (column 3) x total OH divided by 1000.

6. Maintenance factor and/or repair factor - A maintenance factor represents the additional discrepant components (above quantity listed in column 5) resulting from maintenance or handling induced actions. It is listed in percent. A repair factor represents the number of repairable elements of the total failures and is also measured in percent.

7. Number to repair - A product of the total failures x the repair factor.

8. Number of parts required - As a line entry with a component, it indicates the number of end items which must be obtained. It also represents the difference between total failures and number to repair. As a line entry with repair parts, it indicates the number of bits and pieces which must be procured and is the product of the average number of parts required for a repair x the number to repair.

9. Stock on hand - Indicates number of parts held in inventory.

10. Number of parts to buy - This number represents the difference between number of parts required (column 8) and stock on hand (column 10).

11. Quantity of initial spares - Determined by the formula failure rate/operating hour x a constant (usually a value greater than the component maintenance turn around time x operating hours/month).

12. Unit price - Spare part cost per item.

13. Replenishment cost - The product of the number of parts to buy (column 10) x the unit price (column 12).

14. Initial spares cost - The product of the initial spares quantity (column 11) x unit cost (column 12).

15. Holding (storage) cost - This number, provided by the user, represents the costs generated by holding items in the supply systems. This should be provided as a percent of the initial spares cost (column 14).

16. Procurement cost - This is a user provided number. It should be provided in the form of $/item/year.

17. Requisition cost - This is a user provided number. The cost to be entered would be the product of number of parts required (column 8) x $X per item for each requisition.

Cost Worksheet - Labor and Support

1. Part number) - Same as descriptions listed for Cost Worksheet - Material
2. Part name)

18. Maintenance level - Defines the lowest level of maintenance at which the corresponding task in column 22 can be performed. It will be noted as either an operating site or at manufacturer's plant.

19. Maintenance action - This column subdivided into two parts defines the maintenance task in one part and the task time to perform that task in man-hours in the other.

20. Preventive maintenance man-hours - The total time spent for all scheduled inspections on the indicated component. It is the product of the tasks per operating hour x the frequency of the tasks per operating hour x the number of operating hours x the task time in man-hours (column 22).

21. Corrective maintenance man-hours - The total time expended for corrective maintenance on the indicated part. The product of total malfunctions (column 5) or number to repair (column 7) depending on whether a remove/replace or repair task is specified in maintenance actions (column 19) x the task time (column 19).

22. Total man-hours - The sum of preventive man-hours (column 20) and corrective man-hours (column 21).

23. Labor cost - The product of total man-hours (column 22) x the dollar value of 1 man-hour. This man-hour dollar value is user supplied.

24. Tool name - Support equipment nomenclature.

25. Tool part number - Manufacturer's part number.

26. Quantity - The number of each item to be procured considering wearout, breakage, and actual use required to support the system during the life cycle.

27. Unit cost - Cost of tool to customer.

28. Total special support equipment cost - The product of quantity (column 26) x unit cost (column 27).

Cost - Summary

This worksheet summarizes the entire study for each configuration. It presents the subtotal of each of the main support elements within the study and provides, also, the total support cost.

Listed on this sheet are the ground rules used to conduct the study and the assumptions made by the analyst during the study.

INPUT DATA

From Design Engineering

- Identification
  - Part Number
  - Part Name
- Reliability
  - Quantity of Part/Assembly Required
  - Failure Rate/1,000 Hours
- Repair Data
  - Maintenance and/or Repair Factors
- Spares Data
  - Quantity of Initial Spare Parts
- Maintenance Labor
  - Maintenance Task
  - Maintenance Time
- Support Equipment
  - Equipment Identification
  - Quantity Required
  - Equipment Cost

From User

- Quantity Required
- Operating Hours
- Stock on Hand
- Holding (storage) Cost
- Procurement Cost
- Requisition Cost
- Labor Costs

Figure 1

SYSTEM COST SUMMARY WORKSHEET

PROJECT ____________ SYSTEM TITLE ____________

Summary of Support Cost Studies							Facilities		Operating Personnel		Fuel and Utilities		Total Cost
Study No.	Study Title	Selected Config No.	Man-Hours	MMH/OH	Total Failures Per 1000 OH	Cost	Description	Cost	Description and Calculation	Cost	Description and Calculation	Cost	7 + 9 + 11 + 13
1	2	3	4	5	6	7	8	9	10	11	12	13	

OH = operating hours

Figure 3

COST WORKSHEET - MATERIAL

STUDY NO. ______________ STUDY TITLE: ______________

IDENTIFICATION		RELIABILITY			REPAIRS		SPARE AND REPAIR PARTS					SUPPLY COSTS				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Part No.	Part Name	Qty Per Sys	Failure Rate Per 1000	Total OH 1000 x 3 x 4	Maint and/or Repair Factors	No. To Repair	No. Parts Reqd	Stock On Hand	No. Of Parts to Buy	Qty Initial Spares	Unit Price	Replenishment Cost	Initial Spares Cost	Holding Cost	Procurement Cost	Reqn Cost
						5 x 6	5-7		8-9			10 x 12	11 x 12	_% of 14		$__ x 8

COST WORKSHEET - LABOR AND SUPPORT

STUDY NO. ______________ STUDY TITLE: ______________.

IDENTIFICATION		MAINTENANCE MAN-HOUR LABOR COSTS							SUPPORT EQUIPMENT				
1	2	18	19		20	21	22	23	24	25	26	27	28
Part No.	Part Name	Maintenance Location	Maintenance Action		Preventive Maint M-H	Corrective Maint M-H	Total M-H	Labor Costs	Tool Name	Tool No.	Qty	Unit Cost	Total Cost
			Task	Task Time (M-H)	OH Insp x 19 Interval	5 x 19 and/or 7 x 19	20 + 21	22 x $__ hr					26 x 27

COST SUMMARY

STUDY NO. ______________ STUDY TITLE: ______________

Reference Column (From Worksheets) / Support Configuration Concept	SUPPLY					MAINT-LABOR	SUPPORT EQUIPMENT	TOTAL
	Replenishment Cost	Initial Spares Cost	Holding Costs	Procurement Costs	Requition Costs	Labor Costs	SE Cost	Support Cost
	13	14	15	16	17	23	28	

Ground Rules (User supplied)	Assumptions (Designer supplied)
1.	1.

OH = Operating Hours

Figure 2

EXAMPLE 1

Comparing three different design concepts that could satisfy the design requirement.

Ground Rules

100 assemblies
50 operating hours/assembly/month
5 years of operation

Configuration No.	Unit Price	Maintenance Labor Cost	Support Equipment Cost	Other Support Costs	Total Support Cost	Total Maintenance Man-Hours	MH/OH	System Total Failure Rate
1	$13,100	$374,694	$40,560	$6,582,018	$6,997,272	62,449	.238	.01238/OH
2	13,350	335,772	37,560	5,957,502	6,330,834	55,962	.214	.01229/OH
3	13,225	328,794	32,425	5,278,735	5,650,004	54,799	.219	.00883/OH

Total acquisition cost plus support cost must be compared

Configuration 1 - $8,307,272

Configuration 2 - $7,665,843

Configuration 3 - $6,972,504

Figure 4

EXAMPLE 2

Comparing three different maintenance concepts for one design configuration.

Ground Rules

Unit cost $9,327
214 assemblies
50 operating hours/assembly/month
10 years of operation

$\overline{M}$ Concept No.	Maintenance Labor Cost	Support Equipment Cost	Other Support Costs	Total Support Cost	Total Maintenance Man-Hours	MH/OH	System Total Failure Rate
1	$83,416	$2,586,413	$2,034,758	$4,709,587	14,736	.0157	.00108/OH
2	87,018	786,301	2,159,737	3,033,056	14,503	.0154	.00108/OH
3	74,124	93,938	2,959,369	3,126,431	12,354	.0131	.00108/OH

$\overline{M}$ Concept 1 - Major repairs in field

$\overline{M}$ Concept 2 - Major repairs at factory

$\overline{M}$ Concept 3 - No major repairs - throw away and replace

Figure 5

INDEX SERIAL NUMBER — 1112

THE PUBLIC AND TECHNOLOGY: TECHNOLOGY ASSESSMENT

Graham Hereford
Professor of Humanities
University of Virginia and
The Federal Executive Institute
Charlottesville, Virginia
June 1972

Summary

Early in 1972, the House of Representatives of the U. S. Congress passed a bill authorizing an "Office of Technology Assessment", the purpose of which would be to serve the Congress' need for securing competent information on physical, economic, social, and political effects of the applications of technology. The office would "provide an early warning of the probable impacts, positive and negative, of the applications of technology", in order to "assist the Congress in determining relative priorities of programs before it."

The purpose of this paper is to discuss the rationale of technology assessment and to bring out some of the impacts on the work of engineers and scientists, in industry, government, and education. The paper describes the process of technology assessment and estimates some of the economic and other value readjustments it demands.

The object of making a technology assessment is, principally, to describe and evaluate the possible consequences, in the human environment, of developing any particular technology. While such efforts are not altogether new, the requirement for full concentration on all value impacts is novel. More-or-less formal technology assessment is already a legal function of government agencies and their contractors engaged in enterprises that threaten the natural environment. The requirement for assessing natural environmental impacts of technological developments is laid upon agencies of government by the Natural Environmental Protection Act of 1969, specifically in section 102(2) of that Act. As a result of this act, some 2000 "Environmental Impact Statements" had been produced by the end of 1971. Some examples of major assessments: The Trans-Alaska Pipeline, The Amchitka "Cannikin" Project, Metric America, and Snopack.

In spite of early resistance to the NEPA requirements in many quarters, a growing consensus holds that such assessments need to be a regular feature of developing technology. To install this new feature requires that we alter the economic system to include costs of assessment and to reflect environmental and other human social values for which we have not previously accounted.

The basic values to be served by careful technology assessment are concerned with our social and political freedoms. We are coming to realize that while technology has increased our liberties, it may also constrain them in unpredicted and unwanted ways.

In making a technology assessment it is therefore important to consider (among other things) the whole panoply of social values in our country and to see how a proposed technology will alter that set. In addition, alternatives that will strengthen the "higher" values or minimize the impact to the status quo should be addressed.

Because we have a plurality of value sets and systems, and because technological forecasting is a very uncertain process, we need to include, in our assessments, public examinations of possible consequences. The inclusion of public interest reviews as a regular feature of developing technologies is relatively new. It will have important consequences for the engineer and scientist, making them yet more professional in the sense that their responsibilities include the broadest possible knowledge of social consequences.

In 1967, the active Congressman from Connecticut, Emilio Q. Daddario, coined a new phrase -- "technology Assessment" -- as a bid to create a new and important factor in the operations of applied science and technology. In trying to institute a new activity, Congressman Daddario was reflecting long-experienced problems with legislative proposals involving Federal support for developing technologies. In Daddario's words:

> ". . . we have come to realize that the root causes of these problems stem from our inabilities to evaluate long-term world-wide consequences of technology."[1]

The problem of deciding how to avoid ill effects and to produce and enhance good effects had long been complicated, for public officials, by myths about science and technology, by misunderstandings between non-technologist politicians on one side and scientists or technologists on the other, and by narrow views of the social impacts of their work by members of the scientific community.

What was needed, it appeared, was some means of getting a better grip on the possible outcomes of certain technologies that were candidates for strong support. A fair body of literature was beginning, in 1967, to pile up on what we might call major technological mistakes -- very expensive projects that had brought little or no benefit. There were also critical "flaps," like that on ADX-2 that seemed stymied or inconclusive because possible social consequences were never cleared up. Mr. Daddario and his colleagues in the House Committee on Science and Astronautics set out to see if they couldn't discover or develop a capability to anticipate social consequences of possible technology. Hearings were held and draft legislation was developed to serve as a basis for discussion. Finally, the House passed a bill this past February (1972) that would authorize a beginning in a formal way" an "Office of Technology Assessment" to serve the congressional need for "securing competent, unbiased information concerning the effects -- physical, economic, social, and political, of the applications of technology . . . to provide an early warning of the probable impacts, positive and negative, of the applications of technology and to develop other coordinate information which may assist the Congress in determining the relative priorities of programs before it."[2]

Since the 90th Congress, the interest in Technology Assessment has spread widely within government

and a number of academics have tried to get such a process in their sights -- to understand what it entailed and what might realistically and reasonably be its outcome.

What I propose to do in this paper is to expose the scope of this movement as succinctly as possible, and to express some thoughts on the impacts this activity may have on engineers and engineering educators. I would hope to lay out the more important reasons for this development, to describe the process, and to give an accurate general picture of the present and future of such activities. I think it important to explore technology assessment now, in its infancy, for I think it likely it will alter the requirements laid upon us as engineers and as educators.

We can achieve a somewhat fuller grasp on the problem to which technology assessment is addressed by reviewing a few sentences from a report of the National Academy of Sciences.

> "The problems to which we must address ourselves are these: How can we in the United States best begin the awesomely difficult task of altering present evaluative and decision-making processes so that private and public choices bearing on the ways in which technologies develop and fit into societies will reflect a greater sensitivity to the total systems effects of such choices on the human environment. How can we best increase the likelihood that such decisions (domestically, and in the end globally) will be informed by a more complete understanding of their secondary and tertiary consequences and will be made on the basis of criteria that take such consequences into account in a timelier and more systematic way? And how can we do these things without denying ourselves the benefits that continuing technological progress has to offer, especially to the less-favored of the human population?"[3]

Thus one writer on technology assessment defines the enterprise as "the systematic fore-casting, identification, and evaluation of impacts, both beneficial and detrimental, of a technological application within a social context."[4]

Another, less formal definition might be: Technology assessment is an activity that describes and evaluates the possible consequences in the human environment of developing any particular technology. The focus here is on trying to assess what human values a possible technology would have if developed. Of course assessments something like this have been performed ever since there has been any technology, but seldom in a very intense search, and perhaps not often with much objective appreciation of possible alternatives. Also, value impacts have been narrowly construed as economic impacts only. This point needs stress and analysis.

We are accustomed to listening while industrialists complain that recent engineering graduates do not seem to know the importance of the economic constraints on technological development. Now, we are being asked from another quarter to realize either that the constraints are more than economic or that we need to broaden our concepts of economics quite considerably. As a nation we are now trying to assign economic values to factors other than those conventionally deemed economic in the usual economic process. This process of value reorientation is part of the background of technology assessment and provides the reason why technology assessment is more than an academic exercise. The restructuring of the economic rewards and penalities now in process will work to determine which technologies can be economically developed. How long this revaluing process might take to complete will depend upon many factors, but there is no doubt that the process is under way.

At the level of Federal Government, several new agencies are at work altering the economic value structure through new legal constraints. Perhaps the most significant of these are the Environmental Protection Agency, and the Council on Environmental Quality, both created by the National Environmental Protection Act of 1969. This act, in Sec. 102, a, b, and c, instituted what amounts to a requirement for technology assessment. It required that any Federal agency or Federal contractor prepare a statement of estimated natural environmental effects of any activity or project which may possibly have impacts on the ecosystem. Briefly, Sec. 102 of NEPA (1969) requires an interdisciplinary approach involving the natural and social sciences to identify environmental emenities and values that may be affected and to develop a set of alternative action possibilities that will minimize or eliminate adverse effects and enhance positive values.

As of the end of 1971, Federal agencies and their contractors had written some 2,000 of these environmental impact statements. These ranged from a few pages to several thousand, from cursory one-man-month studies to very expensive task-force enterprises taking a year or more. Some famous "Sec. 102" assessments have included those on Calvert Cliffs, Amchitka, and the Trans-Alaskan Pipeline. The latter, published in March of this year presumably provided Secretary of Interior Rogers Morton the basis for approving the North-Slope to Valdez pipeline route in May.[5] The assessment cost about $30 million (according to latest estimates) and occupied more than a year. A couple of examples of other important assessments would be "Metric America," performed by the National Bureau of Standards, leading to the decision that the U. S. should convert to the metric system of measurement, and "SnoPack," conducted by Stanford University and Interior's Bureau of Reclamation to discover and assess the impacts of increasing high mountain area snowfall through winter cloud-seeding.

The last assessment is most interesting. The primary and desired effect would be to increase water supplies for the intermountain region's agriculture. But possibly unwanted secondary effects might dangerously maroon high-mountain residents, induce "cabin fever," and kill their stock. On the other hand, ski resorts might be yet more prosperous -- but at the expense of forested watershed slopes cut down for yet more ski runs. "SnoPack" is thus a fascinating balancing act, in which one seeks a way to justify the wanted primary consequence among ever-widening circles of secondary and tertiary impacts. In the end, the measurable economic benefits might turn out to exceed the costs for some of the population, while other value sets might be affected quite negatively; and these latter might tip the balance. The cost/benefit calculus must focus on the human environmental impacts, and thus finally assess the social impacts directly without relying upon the economic value system to do the work for us. That is the message of the Natural Environmental Protection Act. Slowly but surely the message is getting across, in spite of very adverse feelings in a large segment of the scientific, technological, and industrial community.

The reason for progress in adopting the principles of environmental impact assessment is that the heart

of the message is the rationale of individual freedom.[6] Freedom is a basic value in our system -- a value that we seek to serve through the political process generally. The great popular interest in and support for technology stems from the very real liberties technological development brings to people. Thus, for decades there has been little conflict between the scientific-technological community and the political representatives of the people. Politicians have been quick to support science and technology to improve, protect, and expand the structure of our liberties, in substantial (as well as in trivial) ways. Science and technology have sometimes been playthings but often, also, instruments of great power in the political process.

Technologies both free us and constrain us -- they shape our possibilities. Thus, the question of the shape of our technologies. We are beginning to realize that as our technologies grow in power and number, their effects begin to interact, each with others and all with natural environmental processes. And the effect of all these interactions is to constrain our liberties rather than increase them. That is the political reason for the Natural Environmental Protection Act and for the growing political desire for assessments of the impacts of technology. Technology has given man room and comfort and it has insulated him from the less-tolerable energy interactions with nature. But all this has come at the expense of heavy intervention in the natural system. As human culture has altered and has grown more complex with the onset of each new technology the natural world has grown simpler, in its own terms, and weaker -- less able to complete it cycles as man has more and more invaded.

Man's activities have been so energy-intensive and material-concentrative that natural processes cannot integrate them in the natural evolutionary time span, and process them, as it were, creatively. We have, in short, overwhelmed nature, and the system is in places so degenerated that its biological gains over non-organic processes are failing therefore, to sustain our own and other animal life processes. Thus, air pollution, water pollution, and agricultural failures have proliferated.

As technological applications advance at the expense of the natural environment (and, of course, not all do) they also touch our freedoms and begin to constrain them more or less sharply. It is that impact of technology that needs detecting. Once detected, technological alternatives that continue to enhance freedom may then be imagined and worked out. At least that is the hope the drives technology assessment as an enterprise now considered essential to the political process. For the time being we cannot rely on the economic process as it is now instituted.

Vary T. Coates, in her paper, Examples of Technology Assessments for the Federal Government speaks of the fear that some have that technology assessment may tend to dampen creativity of new applied science. "Yet," she says, "consensus seems to be building that we can no longer depend for social protection on the 'incremental tyranny' of marketing decisions in the private sector nor on piecemeal, *ad hoc* policies of control, regulation, or subsidy by the Federal Government."

Thus, there are new assumptions that are being brought to bear on technological development through the technology assessment process. One of these is that the public has an inherent right to make qualifying inputs toward decisions on developing technologies, that 'naturally' occurring market values reflect only a quite simple and dominant value set and that this set needs to be modified through conscious choices to reflect the natural and human value interaction at the point of technological application.

Another assumption is that since technological developments occur through a temporal series of discrete stages, we may well initiate a development in ignorance of its final form, and, not knowing its final form, we are ignorant of its final cultural, biological, and physical impacts. It behooves us, therefore, to try to predict these impacts and to decide whether we want those impacts.

We are considering technological impacts on people in a pluralist society. Through deeply held values of wide variety, people will react differently to new technologies. Often these reactions will be unpredictable or at least somewhat surprising. For this reason, popular opinion will need to be tested, and public hearings or other methods of public sharing will need to be employed. The Congressional hearing, with all its faults, remains the best test of popular will and sentiment as focused on specific issues. In a democratic society one skips this step at great peril, because through long custom the people feel strongly their right to participate in decisions that will affect them intimately. The tenure (in two-year cycles) of a U. S. Congressman is most sensitive to any slight of constituent rights, as it was designed to be! U. S. Congressmen are very sensitive barometers of the public interest, and in the end it will be their task to formulate technology assessment decisions insofar as those technologies inpinge on the public interest. The political process does not, however, engender complex and novel ideas of the sort that technologists and scientists express. Thus the political process can only act as a check on the creative process, which originates out of the public's sight.

In times past engineers and scientists have accepted and enjoyed their creative roles and their public-servant roles with some innocence. Now, however, being a creative public servant means being judged, in part, in terms of the public value impacts one's work may have. Engineers and scientists (especially in industry) who have found themselves out of the public eye and out of the political arena will more and more be thrust, blinking, into the limelight, insofar as they are truly innovative.

The process of technology assessment will tend to redirect interests of engineers toward more public issues and will end the fairly severe isolation from public policy processes which they have enjoyed (or suffered). In other ways, too, the engineering profession has tended to be insulated from public affairs. Scientifically trained people do not take well or easily to the adversarial style as a means of handling disputes. And yet the technology assessment process involves the conscious attempt to seek out alternatives and to debate the value systems that are touched and furthered or negated by these alternatives. Indeed, more social or political sophistication will be required of those who must engage in the assessment process. And that will involve engineers in much greater degree than it has in the past. Also, more sophisticated economics will be a need. Furthermore there must be some patching or blending of the two cultures that are classically attributed to scientists and engineers on one hand and to humanists on the other.

One not-so-pleasant result of the onset of technology assessment as a public operation will be that certain technical institutions will be open to social

criticism in ways that one would never have dreamed of until recently. For example, such highly respectable institutions as NASA and the AEC have been subjected to considerable criticism from social points of view in the last several years. Both of these agencies have weathered heavy storms of negative and abusive critiques coming from non-scientists and non-engineers.

It used to be possible to pass off such criticism as simply the unworthy or unsophisticated judgments of those who had no right to speak. However, it is now clear that, since technology is having social impacts, some social authorities do have a role in speaking out concerning the acts and judgments of technologists in large organizations, such as the AEC or NASA. We could enumerate and describe other ways in which technology assessment might provide impacts to the education and to work-role development of engineers. But one might summarize by saying that after many long years of constant reiteration of need for engineers to become professionals, it now looks as if professionalism is truly being thrust upon engineers through the gradually growing recognition that the social meaning of technology has high intensity with a high order of complexity; and insofar as we can become aware of the social impacts of the work of engineers, we stand responsible for those social impacts in ways that are similar to the responsibilities of attorneys and medical doctors. The engineer stands very near the attorney in social importance as one of the architects and shapers of modern, post-industrial society. The society should be called post-industrial precisely because the industrial ideology is now passe; and we are aware of and responsible for the values of our technological development in ways quite beyond awareness in the earlier, industrial era.

Footnotes

1. Technology Assessment--1970, Hearings before the Sub-Committee on Science Research and Development, Part II, p. 340.

2. H. R. 18469: "A Bill to Establish an Office of Technology Assessment for the Congress," September 9, 1970, p. 3.

3. National Academy of Sciences, 91st Congress, 2nd Session, "Technology: Process of Assessment and Choice," House Committee on Science and Astronautics, Committee Print, 1969."

4. Vary T. Coates, The Federal Government and the Current Development of Technology Assessment, Occasional Paper No. 11, June 1971, The George Washington University, Program of Policy Studies in Science and Technology, p. 1.

5. The fact that this decision is now undergoing court tests and may be reviewed by the U. S. Supreme Court shows clearly the strong conflict of new political values with old technological and business values.

6. The suggestion of a connection between technological development and political freedom is due to Franklin P. Huddle of the Congressional Research Service of the Library of Congress. Frank Huddle has written extensively and well on technology assessment. See, for example, "Government Technology Assessment: The Role of the Social Sciences," The Library of Congress; Legislative Reference Service (Washington, D. C., October 2, 1970).

THE EVALUATION OF FAILURE AND FAILURE RELATED DATA

W. E. Vesely
Aerojet Nuclear Company
Idaho Falls, Idaho 83401

SUMMARY

Techniques for evaluating failure data are described which yield a wide range of potentially useful results. From basic data sources such as failure logs, the evaluations can yield not only failure rates, but can uncover wear-out and degradations, environments causing unwarranted failures, can assess repair and maintenance effectiveness and can locate sub-par performers. The particular techniques involved in these evaluations are discussed and data requirements and practical applications are given. Computerization is also described. The aim of the paper is somewhat elementary, merely to give the requirements and potential returns of data evaluation so that it can better be decided whether such an undertaking is actually and not only potentially worthwhile.

I. INTRODUCTION

In the general field of failure reporting, a wealth of data exists pertaining to failures of components, subsystems and systems. The reported failures include minor and major accidents, safety and reliability failures, and electrical and mechanical equipment failures. Even though the data exists, it is rarely evaluated to extract quantitative results which objectively characterize not only the failure, but the causes of the failure. If any evaluation is performed, then failure rates are usually the only information extracted. This, however, represents only a fraction of the information obtainable. A more complete list of quantitative results which can be yielded by evaluation includes:

1. The obtainment of component and system failure rates.
2. The determination of abnormal environments and stresses causing unwarranted failures.
3. The assessment of testing and repair effectiveness including the detection of unwarranted burn-in failures.
4. The detection of degradations in component and system performance including the onset of wear-out.
5. The evaluation of maintenance effectiveness.
6. The identification of components or systems which are deviating in performance from their peers.

These results can be obtained from any data which records repetitive-type failures, that is, the failures can be repaired or if they are nonrepairable, then the components or systems failing have other identical or similar counterparts in existence.

The above extractable information offers the reliability or safety engineer a means of auditing, predicting and correcting failures. In particular, the extractable information offers the following potentials:

1. A means of objectively evaluating system performance with regard to reliability and safety.
2. A means of objectively assessing the efficiency of a maintenance and repair program.
3. A means by which deficiencies in existing systems can be pinpointed and corrected - early and before the deficiencies actually cause system failure.
4. A means of quantitatively assessing the impacts of design modifications and maintenance changes.
5. Finally, a means by which realistic data can be obtained to be used to predict future reliability and safety performance.

These potentials are extremely practical and hence extremely attractive since they pertain to an existing, real life situation; the evaluations are not done on some mathematical model, but are done on a physical system or phenomenon in order to yield physically applicable results.

It must be recognized of course that these potentials represent an "optimum" which in practice is never achieved. In practice, there are definite obstacles such as the problem of separating true failures from routine maintenance calls and the problem of identifying the true causes and not merely the symptoms. The potentials however are there and seem worthwhile enough to attempt even partial fulfillment. The true test comes of course in their payoff in economic returns.

This discussion does not pretend to analyze the true returns from these potentials. It will however attempt to describe the types of evaluations involved in theoretically achieving these potentials. The emphasis will not be on the mathematics, but on the applications and interpretations. Practical implementation of the techniques will also be described. With this discussion of what is involved in the potentials, it is the hope of this paper that one can then better decide whether such an undertaking is worthwhile and whether the true return exceeds the investment.

II. THE DATA INFORMATION REQUIRED FOR QUANTITATIVE EVALUATIONS

The basic piece of data needed for a quantitative evaluation is the time of the occurrence of the failure.[a] The precise minute and hour of the failure is not necessary and, in fact, is generally useless for the previously described evaluations. The day of the failure is usually the finest resolution needed and is only needed when failures occur frequently (on the order of once a week). A favorable circumstance in data recording is that the resolution needed for the time of the failure decreases as the number of failures decrease; for those failures which occur on the order of once a month the approximate day (plus or minus several days) is needed and for those failures which occur less frequently the week or month of the failure is only needed.

In addition to the time of the failure, the only other basic data needed is the identification of the failure. This identification should optimally be a categorized identification with succeedingly finer resolution given in the lower category levels. The general nature of a categorized identification is given below:

1. The system in which the failure occurred.

[a] Here, and for the rest of the report, a "failure" means any "abnormality", covering the spectrum from a consequential failure to merely minor trouble.

2. The subsystem(s) of the failure occurrence.
3. The component suffering the failure.
4. The criticality and mode of failure.

For finer resolution of identification, the above categories would be subdivided into sublevels and for simple accessing, the categories would be keyed by some indexing scheme.

The categorized identification of the failures is optimal since it allows various detail in the evaluation, from a system to a mode of failure examination. However, any identifiers of the failure will yield the same type of results as listed in the INTRODUCTION. The resolution of the identification does not affect the nature of the results obtained nor does it affect the amount of information obtained. The resolution only affects the discriminating ability of the results, i.e., their "fineness".

III. EVALUATIONS AND THE EXTRACTED RESULTS

The following sections briefly describe the types of evaluations which are used to obtain the extracted results listed earlier. Incorporated within certain of the evaluations are statistical tests which screen the noise from the physically significant behaviors, and the nature of these tests is also described. In all of the sections, the emphasis will not be on the mathematical details, but will be on the applications of the evaluations.

A. The Obtainment of Failure Rates

The failure rate is the basic parameter which characterizes the reliability or safety of a unit (or incident). If λ is the failure rate and R is the probability that the unit will suffer no failure to time t, then R is simply given by the relation

$$R = e^{-\lambda t}. \qquad (1)$$

In reliability engineering, the quantity R is termed the reliability and is simply the percentage of time the unit will suffer no failure when it is operated for a period of t hours (or days). Where the failure is not associated with a unit, such as an incident, R simply gives the probability that the incident will not occur in the time period t.

The failure rate is related to the mean time between failure T by the equation,

$$\lambda = \frac{1}{T}. \qquad (2)$$

The quantity T is simply an average of the times between failure occurrences. One may start at any failure for this averaging, and hence, the installation date is not necessary when the repair or replacement time is small compared to T.

The above two equations serve as the basis for determining the failure rate from the failure data recordings.[b] A complete statistical treatment shows that in order to determine λ any one of three types of data can be used:

1. the number of failure occurrences,
2. the time of the last failure, or
3. the times between the failures.

[b] As stated earlier, the failure data recordings can be in the form of failure reports, maintenance histories, trouble call recordings, etc.

Corresponding to the three types of data used, three methods exist by which the failure rate can be determined. Each of the above methods has its own merits and advantages depending upon the particular circumstances; however, any one is sufficient to yield the failure rates. The capability of determining the failure rate in a number of ways offers the advantage of the evaluation being able to be adapted to any peculiar data.

Table 1 depicts the type of failure rate printout which is yielded by a straightforward computer program. In this particular instance, the mean time between failure was output (the failure rate is simply the inverse of the mean time between failure). See Table 1 at the end of this text.

In Table 1, the "Group Number" and "Unit Number" denote the particular indexing used in the failure record. The group number identifies the subsystem and the unit number identifies the component within the subsystem. The first three columns give the mean time between failure for the corresponding component for 1969, 1970, and for the first three months of 1971. The column denoted by "2.25-Year" gives the average mean time between failure for this 2.25 year period. The last three columns give the 90% confidence bounds (error bounds) for the mean time between failures. (The error bounds for 1969 were not requested in the computer printout.) A dash in an error bounds column denotes the fact that no failures occurred in the particular period and hence no upper bound could be obtained.

B. The Determination of Abnormal Environments and Stresses

The statistical test described in this section determines any failures which are not occurring randomly. From basic reliability and safety theory, if failures occur randomly, then the times between failures follow an exponential distribution. These random failures are to be expected and have no prescribable cause as to their occurrence. If the failures are not random and are due to some physical cause, then departures from the exponential will be observed. For example, if the failures are due to some abnormal environment condition (such as heat build-up), then the times between failure will show peak characteristics in its distribution.

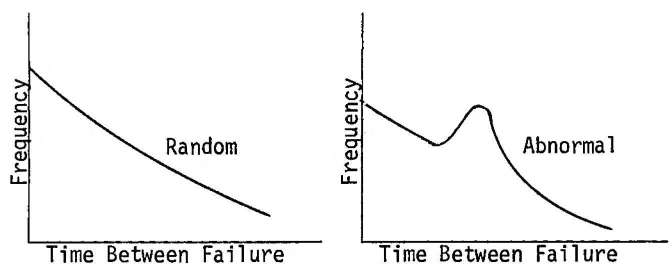


Figure 1. Failure Frequency Versus Time Between Failure For Random And Caused Failures

For the statistical test, termed an "exponential test", which is a simple chi-squared test, times between failures are needed as the only data. On the average, ten failures are needed in order to determine abnormal causes for the failures. However, for more extreme abnormalities, four failures suffice to show these causes. The statistical test, which is essentially a rejection-type test, screens the expected perturbations from the unexpected, caused perturbations.

In practical application, the exponential test is used to determine abnormal conditions which exist and which cause unwarranted failures. Stresses within the

component or system, nondesigned operating environments, and in general, any abnormal impressed condition can be uncovered by the test. The test is also useful in determining whether a failure rate (λ) validly defines the failure occurrences; for these failures which follow an exponential distribution, the failure rate can be used in further applications as discussed in the previous section. For those components that normally follow nonexponential distributions, (such as Weibull behavior) the particular distribution may be used as the base in determining abnormal behavior.

Table 2 shows the results of the exponential test applied to physical failure data. The results were extracted from the output of a computer program which automatically tests all the data in the failure records. In the table, the "Group Number" and "Unit Number" identify the subsystem and the particular component within the subsystem. Thus, an abnormality condition was detected for Component 46 in Subsystem 39. The departure from an exponential distribution was significant at the 90% confidence level and the "Total Deviation" (6.3) gives the magnitude of the departure (no departure is zero deviation). This deviation number can be used to further rank the abnormalities obtained. In the computer program, the times between failure are grouped into various intervals and the "End of the Intervals" column gives the interval limits. The dash symbol in the last column denotes that the limit of the last interval is infinity. See Table 2 at the end of this text.

Comparison with similar type components showed that this was not a normal type failure behavior. The exponential test consequently showed that this particular component was experiencing an environment which caused it to depart from a random failure behavior. The environment incurred a peaking in the interval from 45.4 days to 67.1 days, causing four times more failures than expected. Correcting this environment maximally could reduce the number of failures occurring in the interval by a factor of four which in turn would reduce the total number of failures occurring by 40%.

C. The Assessment of Repair Effectiveness

The exponential test of the previous section can be simply extended to determine ineffective repairs and unwarranted burn-in failures. From basic precepts of renewal theory, if the repair or replacement is effective, then the time between failure distribution will be exponential. Similarly, an exponential distribution will result if excessive burn-in failures ("sudden deaths") are not being experienced. For ineffective repair or excessive burn-in, on the other hand, the distribution will exhibit a peak for small values of the times between failure.

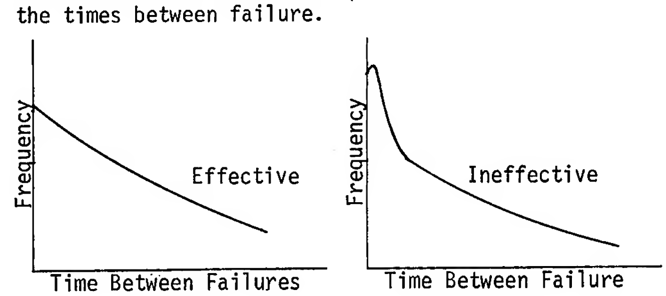


Figure 2. Failure Frequency Versus Time Between Failure For Effective And Ineffective Repair

If the failures are repaired when they occur, this initial peak in the distribution simply means that one is having to come back too soon to repair the item again because it was not effectively repaired the first time. This ineffective repair may be due not particularly to actual repair inefficiency, but to undiagnosed troubles or may be due to hardware effects such as poor quality parts, unusually severe usage, etc. If the items which fail are replaced instead of repaired, then the initial high values in the distribution depict a high percentage of burn-in failures occurring, such as when the items have not undergone burn-in within the initial quality control program. In either case, repair or replacement, the large number of failures experienced within short intervals denotes an ineffectiveness in operation which can often times be corrected.[c]

When a repair defect is uncovered, the particular operational circumstances may dictate that this ineffectiveness cannot be corrected and is a fact of life. This is the case, for example, when repair is very difficult and several attempts are needed (such as O-ring alignments). It is also the case when burn-ins must be tolerated (e.g., when destructive testing is the only means of investigation). The test results will simply point out these difficulties. In a number of other cases, however, the ineffectiveness can be corrected to yield returns, particularly since the correction will significantly increase the mean time between failure.

Table 3 shows the results of the test when applied. The evaluation was part of the automated program described in the previous section. For this failure record, a three day interval length was chosen by the program as the maximum time between failure value (the ineffective repair indicator), and the failures occurring within three days apart were extracted from the record. The expected number which should occur within three days was computed and compared to the actual number which had occurred.

Of the total number of failures occurring, 53% had times between failure less than or equal to three days; this meant that in 53% of the repairs, one had to return within three days because the item had failed again. From the data, one should only expect 14% of the repairs having to be repaired again within three days. The test determined that this discrepancy was in fact real and an ineffectiveness existed. Correcting the ineffectiveness would increase the mean time between failure from 19.2 days to 33.2 days. See Table 3 at the end of this text.

D. Evaluations of Maintenance Effectiveness

Statistical evaluations can be performed on the failure data to answer the following questions: "Has a change in maintenance been effective in reducing failures"?, "Have the number of failures significantly increased or decreased in a given period"?, and "Has a design change effectively reduced failures"? Since failure behavior exhibits noise-like characteristics, simple observations cannot always be made to obtain answers to these questions. Even though a unit is maintained at constant performance, with no upgrade or downgrade occurring, it can suffer more failures in one period than in another period. This is true simply because of the random nature of failure behavior. This random-like behavior occurs to even more of an extent when maintenance is being performed on the unit; at times maintenance is better than average and at other times it is worse than average.

Table 4 shows the evaluations (chi-squared and F-tests) applied to the yearly number of failures occurring in a system. The results are extracted from a computer run in which all the failure records were analyzed. The

[c] As in the previous section, instead of an exponential the known normal behavior distribution may be used as the base for comparison.

unit numbers in the table are individual component identifiers for this particular failure record. As observed from the table, the true changes are straightforwardly differentiated from the noise behaviors. For example, Unit 03 failures decreased from three in 1969 to one in 1970, but the test showed that this could not be attributed to any true decrease. No significant change was therefore given to the change in number of Unit 03 failures. In addition to the above failure change evaluations, the program has the capability of also obtaining the dominant failure contributors of a subsystem, or system, i.e., those which are causing more than their share of failures. See Table 4 at the end of this text.

Used in the above surveying-type manner, the test serves as a simple tool for auditing the number of failure occurrences, showing the dominant failure contributors, the significant increases, decreases and steady state performances. The "bad actors" (increases and dominant contributors) are flagged for possible further investigations and the "good actors" (decreases and steady actors) show effective modifications and effective maintenance.

E. The Detection of Upgrades and Downgrades

During the course of operation, certain of the components and systems may suffer a degradation such that their chance of failure increases (for example, components wearing out). These units will subsequently start to fail more often. Where the failure is not hardware associated, but is a general incident occurrence, the degradation may take the form of conditions changing such that the chance of the incident occurring increases. Whether the degradations are hardware associated or incident associated, when such degradations occur, performance and safety decrease and cost expenditures increase. From both an economic and safety point of view, these degradations need be identified and if the degradation is critical, they need be corrected.

For the test, the times of the failure occurrences are used as the basic data. On the order of six failures are needed as minimum data. These failures may have any detail of identification; if the failures are classified only to a subsystem level, then the subsystem performance will be evaluated for a downgrade or upgrade. If the failures are identified to a component mode of failure level, then the performance of this mode of failure will be evaluated. The performance test is thus quite versatile in its range of applicability.

The test used in detecting downgrades or upgrades is a straightforward adaption of the standard statistical F-test. Table 5 shows the results of the F-test applied to failure records. The table was extracted from the output of a computer program which analyzes the failure records of a data source and detects any downgrades or upgrades.

The failure times in the table denote the times of occurrence of the failure, given as the month of the failure (M), the day of the month (D), and the year (Y). The failures are numbered sequentially and these are the left-most numbers. The times between failures (days) are simply the intervals of time between the subsequent failure occurrences. For this component, the test detected a significant downgrade in performance. This downgrade resulted from a comparison of the mean time between failure for the first six failures to that for the remaining five failures, and this is the reason the "downgrade" symbol is beside the sixth failure. For the first six failures, the mean time between failure (T_1) was 71.4 days and for the remaining five failures, it was 18.4 days (T_2). If one would assign a beginning time to the degradation it would thus be between the sixth and seventh failure occurrence. The ratio of change in the table is the ratio of T_1 to T_2. The mean time between failure thus had decreased by a factor of 3.88 which meant that on the average the failures were occurring 3.88 times as often as before. The factor of 3.88 was greater than the noise level value and hence the change was a true degradation. See Table 5 at the end of this text.

F. The Identification of Deviate Performances

The last test described in this paper identifies deviate performances. From a selected group of components or systems, the test will determine those particular units which are behaving differently from the rest. Those units which are significantly worse than the others and those which are significantly better than the others will be obtained. These deviate performers will be identified because of their failing more frequently or less frequently, but will also be identified because of their different failure behavior. The deviate identification due to frequency of failure allows one to locate the good performers and bad performers of the group, those which have better or poorer mean times between failures. The deviate identification due to behavior allows one to investigate the actual distribution of failures, enabling one to determine how units are failing and to determine the causes for their different behavior. The test involves the standard Smirnov test and F-test.

For the "deviate performance test", a group of items are first selected for the comparisons, where the items may be components, subsystems, incident occurrences, etc. Each item in the group, representing one series of failures, is then compared with the other items of the group. The failure times of each item are used as the data in the comparisons. A minimum of two failures is needed for each item and the items compared may each have a different number of failures. As is the general case for all of the tests, the more failures recorded, the finer is the resolution of the test.

Table 6 illustrates the output of a computer program which utilizes the deviate performance comparisons. In this particular instance, comparisons were made of similar components located in the same and in different subsystems. The test was performed to identify any "bad actors", either due to hardware defects or to adverse environments experienced. For each intercomparison, the program cycles through all the quantities to be compared and prints out a simple, legible output as illustrated in the table. See Table 6 at the end of this text.

In the table, analogous to the previous examples, the "Group Number" identifies the subsystem and the "Unit Number" identifies the individual component within the subsystem. The "Individual MTBF" is the individual mean time between failure for the particular component and the "Average MTBF" is the average mean time between failure for the remaining components. The mean time between failure is in units of days. The "MTBF Ratio" is the ratio of the individual to the average.

In the column labeled "MTBF Comparison", the word "worse" means that the component mean time between failure is significantly worse than the rest and "better" means it is significantly better than the rest. No descriptor in the column denotes no significant difference from the rest. The "Distribution Comparison" column has a similar interpretation with no descriptor indicating it has the same distribution.

With the deviate performers identified, the reliability or safety engineer has information on which to make decisions. Action, for example, can now be taken. The components which have mean time between failures which are worse than the rest can be replaced or upgrade action can be taken. One notes that, in general, the Components of Group 1 are worse than those in Group 2 (all the "worse" components are in Group 1 and Group 2 has all "better" components). Hence, a subsystem degradation is indicated.

The different distribution results denote the fact that these different failure behaviors are critical enough to cause a deviate performance. The individual distributions can now be obtained to determine the specific causes for the deviations. As discussed previously, if a high peak exists at small times between failure, ineffective repair is indicated. If the times between failure are becoming successively smaller, a wear-out may be indicated. A number of the previous tests evaluated the data for these types of characteristics and their output can be used to help uncover the causes for deviations.

Table 1. Mean Time Between Failure Results (in days)

GROUP NO. 28

Unit No.	1969	1970	1971 (3 Mo.)	2.25-Year	1970 (90%)	1971 (90%)	2.25 (90%)
06	121.7	45.6	90.0	68.3	(25.3, 91.7)	(19.0, 1754.7)	(42.2, 118.4)
07	121.7	91.3	45.0	91.1	(39.9, 267.1)	(14.3, 253.3)	(52.2, 174.6)
08	121.7	45.6	90.0	68.3	(25.3, 91.7)	(19.0, 1754.7)	(42.2, 118.4)
09	91.3	60.8	90.0	74.5	(30.8, 139.7)	(19.0, 1754.7)	(45.0, 132.9)
10	60.8	33.2	45.0	43.2	(20.0, 59.2)	(14.3, 253.3)	(29.4, 66.0)
11	91.3	36.5	90.0	54.7	(21.5, 67.3)	(19.0, 1754.7)	(33.5, 88.7)
12	365.0	121.7	90.0	273.3	(47.1, 446.4)	(30.0, ------)	(105.8, 1002.8)
13	121.7	36.5	12.9	41.0	(21.5, 67.3)	(6.8, 27.4)	(28.1, 62.0)
14	60.8	73.0	90.0	68.3	(34.7, 185.3)	(19.0, 1754.7)	(42.2, 118.4)
15	365.0	121.7	90.0	205.0	(47.1, 446.4)	(30.0, ------)	(89.6, 600.2)
16	182.5	60.8	90.0	91.1	(30.8, 139.7)	(19.0, 1754.7)	(52.2, 174.6)

Table 2. Exponential Test Results Showing An Abnormality (90% Confidence)

GROUP NO. = 39 UNIT NO. = 46

TOTAL DEVIATION = 1.4 (8 Failures)

End of Interval (days)	Expected Percentage	Observed Percentage	Deviation
8.3	14.3	0.0	0.0
18.0	14.3	14.3	0.0
30.0	14.3	0.0	0.0
45.4	14.3	14.3	0.0
67.1	14.3	57.1	6.3
104.2	14.3	14.3	0.0
-----	14.3	0.0	0.0

Table 3. Test Results Indicating Ineffective Repair or Replacement

(90% Confidence)

Percentage of Failure Occurring Within Three Days Apart

EXPECTED	OBSERVED
14%	53%

Mean Time Between Failure Consequences

EXPECTED	OBSERVED
33.2 Days	19.2 Days

Table 4. Audit of Failure Changes

(To 90% Confidence)

KEY

↓ Significant Decrease In Failures From 1969 to 1970
↑ Significant Increase In Failures From 1969 to 1970
(No Symbol Denotes Steady State Behavior)

Unit No.	No. of Failures In 1969	No. of Failures In 1970	Change From 1969 to 1970
01	5	0	↓
02	5	0	↓
03	3	1	
04	4	1	
05	6	15	↑
06	2	1	
07	0	1	
08	0	4	↑

Table 5. Results of the Performance Test Showing A Downgrade (90% Confidence)

GROUP NO. = 48 UNIT NO. = 66

	Failure Times M/D/Y	Times Between Failures (Days)	Type Of Change*	Ratio Of Change	T_1 (Days)	T_2 (Days)
1	5/19/69					
2	5/28/69	9				
3	6/17/69	20				
4	11/10/69	146				
5	11/12/69	2				
6	5/11/69	180	Downgrade	3.88	71.4	18.4
7	6/07/70	27				
8	7/01/70	24				
9	7/02/70	1				
10	7/07/70	5				
11	8/11/70	35				

*No symbol in the column denotes steady state behavior.

Table 6. Deviate Performance Results
(90% Confidence)

Group No.	Unit No.	Individual MTBF	Average MTBF	MTBF Ratio	MTBF Comparisons	Distribution Comparisons
1	1	48.0	105.1	0.46	Worse	Different
1	3	23.0	104.4	0.22	Worse	Different
1	5	128.4	97.7	1.31		
1	7	92.8	99.3	0.93		
1	9	136.4	97.4	1.40		
1	16	57.4	102.7	0.56	Worse	
1	17	397.0	96.5	4.11	Better	
2	3	92.0	99.4	0.93		
2	10	451.0	96.1	4.69	Better	
2	14	95.8	99.6	0.96		
2	15	114.5	98.7	1.16		
2	18	248.0	96.5	2.57	Better	
2	20	82.0	99.2	0.83		
2	23	181.0	96.3	1.88	Better	

The MTBF Of The Entire Ensemble = 99.1 Days

QUALITATIVE ANALYSIS OF REACTOR PROTECTIVE SYSTEM

B. M. Tashjian, Supervisor
Reliability Analysis
Nuclear Power Department
C-E Combustion Division
Combustion Engineering, Inc.
Windsor, Connecticut

Abstract

The application of the failure modes and effects analysis technique to the reactor protective system is described. A simple example is used to illustrate the method. The procedure for evaluating the system for compliance to the single-failure criteria is highlighted.

Introduction

The design of reactor protective systems must meet certain criteria to make sure that the systems are acceptable for use in operating plants. These criteria specify functional performance requirements for operating environments, and specific design features necessary for the system to be licensable. The basic criteria is published in the Federal Register.[1] Supplementary criteria have been and are continuing to be developed by recognized standards development groups. Examples which will be used in this paper are ANSI N42.7[2] and ANSI N41.14[3].

In ascertaining whether a system complies with these requirements, it becomes necessary for a manufacturer to analytically test his system. One requirement in particular is adherence to the single-failure criterion.[4] This criterion requires that a single failure must not render the protection system incapable of performing its intended function. A procedure for evaluating the system for compliance is the failure modes and effects analysis technique. The questions that are addressed in applying this technique are what single failures can be potential "violators" of the single failure criteria, and which single failures are undetectable?

The Reactor Protective System

The function of the reactor protective system (RPS) is to automatically initiate reactor protective action whenever selected nuclear steam supply system (NSSS) parameters monitored by the system reach a preset level. The protective system is designed in compliance with the IEEE standard: "Criteria for Nuclear Power Plant Protection Systems (ANSI N42.7)." These criteria require the system to meet specific design requirements. The principal ones are:

Automatic Action

The system shall automatically initiate appropriate protective action whenever a condition monitored by the system reaches a preset level with precision and reliability.

Single Failure

Any single failure within the protection system shall not prevent proper protective action.

Channel Independence

Redundant protective signal channels shall be independent and physically separate to accomplish decoupling of the effects of unsafe environmental factors, electrical transients, and physical accidents.

On-line Sensor Check

A means shall be provided for checking operational availability of each system imput sensor during reactor operation.

Test and Calibration

Channels shall be capable of being tested and calibrated either during the station shutdown or during power operation.

Channel Bypass or Removal

The system shall be designed to permit any one channel to be maintained, and when required, tested or calibrated during power operation without initiating a protective action at the system level. During such operation, the active parts of the system shall continue to meet the single-failure criterion.

Manual Initiation

The protective system shall include means for manual initiation of each protective action at the system level. No single failure within the manual, automatic, or common portion shall prevent initiation of protective action by manual or automatic means.

The design of the system used as an example here meets these criteria while maintaining a high degree of plant availability. The system achieves both these objectives by utilizing four independent measurement channels for each NSSS parameter monitored. The NSSS parameters which can initiate reactor protective action are power level, rate-of-change of power, primary coolant flow, steam generator water level and steam pressure, pressurizer pressure, reactor thermal-margin, and loss-of-turbine load.

The protective system also provides for conversion of the system logic for each monitored parameter from a two-out-of-four coincidence to a two-out-of-three coincidence logic requirement for initiation of protective action. This provision allows a single channel of each monitored NSSS parameter to be removed from service for maintenance, test, or calibration while still maintaining a system which can accommodate a single failure in any measured channel without initiating inadvertent protective action.

A technique that will be applied to evaluation of the system for compliance with the design criteria is the failure modes and effects analysis (FMEA) study. This study is an analytical tool which can provide useful information for selecting design alternatives, corrective action priorities, and test planning criteria. A discussion of the FMEA can be found in the general principles for reliability analysis document, ANSI N41.4.

Failure Modes and Effects Analysis

The failure modes and effects analysis is a systematic procedure for analyzing a system from the point of view of component failure. The procedure is used to study the failure modes of the components in the system and to determine their effect on the system performance at local and overall levels. The results of the analysis can be used for assessing the system compliance with the single-failure criteria, and for developing mathematical models for quantitative studies.

As in any analysis, it is necessary that the boundaries of the system be clearly defined and that the level of detail of the analysis be established. The boundaries and definition may vary depending upon the timing of the analysis and the impact the results would have on design changes. These conditions should be stated in order to understand the extent of the analysis and to properly interpret the results of the work.

The components of the system contained within the boundaries are grouped together into an appropriate block diagram. A functional block diagram is made and each block within the system, including its components, is identified. The failure modes for each component are determined, and their effect on the local and overall system levels is analyzed and tested to the single-failure criteria. A probability of each failure mode is assigned. The probability for the overall system can be calculated.

Level of Analysis

The level at which the FMEA was conducted was determined by the functional stratification of the system. The replaceable item level was chosen. It was picked because any changes or additions in the design resulting from the FMEA study would most likely be made at this level and result in the least cost impact. The effects of the component failure modes were also studied at the next higher functional system level and at the overall system level. The replaceable items consist mostly of electronic component parts contained within each block. The generic classification of these parts is shown in Table I. The effects of these parts' failure mode were determined first at the block level and then at the system level. The resulting effects on the system level were evaluated with respect to the intended function of the RPS.

FMEA Worksheets

The analysis was conducted on worksheets (Fig. 1). The FMEA worksheet provides a systematic layout for tabulating information and keeping track of the analysis, the format being designed to serve this purpose. The analysis is conducted by identifying components in the subsystem, listing their failure modes, and studying their effects on system performance. These effects were observed for various operating modes of the system. A description of each entry in the worksheet follows:

1. Diagram -- This diagram entry depicts the functional relationship between each item under investigation in the analysis to one level of detail greater than that shown in Fig. 2. For example, the level of detail in the diagram is to the items which are line replaceable; e.g., sensors, voltage, comparator cards, etc. Failure modes are assigned to the lowest level of detail shown in the diagram. The analysis is then conducted for this level for each of the block outputs.

2. Number -- The number used is for information and tabulation purposes. It serves as a reference indicator in the summary of results.

3. Name -- Each element in the diagram is identified by a name. The analysis is conducted at this equipment level.

4. Failure mode -- All significant failure modes, including both random and degradation failures, of the elements comprising the diagram are evaluated. The occurring failure modes that have been reported on similar items were considered predominant and used in the study. Some of them are:

Item	Failure Mode
Resistors	Open, short, drift
Diodes	Open, short
Capacitors	Open, short
Relays	Open coil, shorted coil, contact fails to transfer

5. Cause -- The most probable cause associated with each failure mode is listed. These causes are generally related to the next lower echelon of equipment breakdown and to the circuit and principle environment parameter sources.

6. Symptoms and local effects -- The immediate consequence of each failure mode, along with a dependent failure or secondary side effects resulting from the possible cause, is determined. These symptoms are generally examined at one equipment level higher than the item which has failed in entry 3.

7. Method of detection -- This entry lists the mechanism within the system which detected or indicated the occurrence of the failure mode. The failure effect could be annunciating or nonannunciating to the operator. If it is nonannunciating, the method of detection includes the means by which the operator can detect the failure such as use of external test equipment, periodic performance checks, etc.

8. Inherent compensating provisions -- This entry

lists the existing circuitry within the block (diagram) that will compensate for the failure mode at the level being analyzed. It excludes redundant circuitry in other parts of the system, unless so identified.

9. Effect upon -- This entry lists the ultimate effect of the failure mode on the next higher level of equipment breakdown than in entry 6.

10. Failure probability -- This is a quantitative value of the occurrence of the particular failure mode of interest; and it may be represented as a relative value.

11. Level of severity -- This is a judgement of the failure effect on the overall system and is used for identification.

12. Remarks and other effects -- This entry lists the effects of this particular failure mode on the overall system performance. Effects which may not be recognized locally, but can be observed on a system level, are entered in this column.

Example of FMEA

The model illustrating the analysis is a simple protection system sensing one parameter, pressure. This model in its simplest form is shown in Fig. 2. The function of this model is simple: pressure level is continuously monitored and when a preset limit is exceeded ($P_M > P_S$), a protective action signal is initiated. The protection action for this particular parameter is the initiation of a reactor trip. The protective action signal interrupts electrical power from a motor-generator set to the drives for the control rods in the reactor. The removal of this electrical power releases the control rods from their motive source and causes them to insert into the reactor core (TRIP).

The notation used in the block in the diagram (Fig. 2) represents the functional logic of the system: namely, two-out-of-four coincidence logic system with four redundant trip paths. Pressure (P_M) is monitored continuously by four independent and identical sensor circuits in the system. When the pressure level exceeds a preset limit (P_S), the sensors detect the level change and initiate measurement channel protective signals. The protective signals pass through the coincidence logic gates and four trip paths to the trip actuation devices completing the protective action.

At least two channel protective actions are needed for reactor trip. The six two-out-of-four system logic matrices provide protective signals to four independent trip paths (1,2,3, and 4 in Fig. 3). The trip paths control the operation of four sets of circuit breakers located between the motor-generator sets and the drives for the control rods. During normal operation, the breakers are closed connecting the control rod drives to their power source. The trip circuit breaker sets are arranged to allow for testing of the system during reactor operation, while maintaining the protective function of the system. Any of the four trip paths can complete the trip function of the system for protective action. The trip paths control the circuit breaker operation between the motor-generator sets and rod control power supplies and cause the breakers to open and remove electrical power from the control rods. The control rods will release and drop into the core completing the protective action.

Functional Block Diagram

The functional block diagram of the system is shown in Fig. 3. There are four redundant sensor channels shown terminating into the coincidence logic block, and four redundant trip paths from the logic block to the trip function block. The function of the system can be described in terms of the function of its constituent blocks and analyzed accordingly. Simply, the system monitors pressure, and initiates protective action when a preset pressure is exceeded. This action must be accomplished satisfying specific performance and design criteria. Those criteria include redundancy, testability, single failure, etc., and are met in part by the function and location of equipment in the system. The equipment is located in functional blocks and can be so identified. These blocks are arranged in a logical manner (Fig. 3) so that for normal operation the system complies with the criteria as well as achieving its objectives -- providing protective action. The description of these blocks and their primary function is given in Table II.

Each sensor channel consists of a sensor power supply, voltage comparator circuit (bistable), and trip relays. The output of the sensor channel is a logic signal. The logic signal is achieved by trip relays. The contacts of the trip relay are located in the coincidence. If the measured pressure level exceeds the preset reference point setting ($P_M > P_S$), the output signal from the sensor channels changes their logic state. This change causes the trip relay coils to deenergize and in doing so, transmits the protective action signal to the four trip path circuits through the coincidence logic block by means of the trip relay contacts.

The two-out-of-four coincidence logic is composed of six two-out-of-two AND gates. Each AND gate requires two input signals for an output. The AND gates are themselves the trip relay contacts arranged in pairs. The contacts are arranged in parallel and require both contacts to open for AND gate operation. Each contact is controlled by one sensor channel with no two contacts from the same channel. In the AB AND gate, then, one contact is controlled by the A channel, and the other by the B channel. The six gates are designated as AB,AC,AD,BC,BD, and CD for the four sensing channels A,B,C, and D. When taken collectively, any two of four channel signals can initiate coincidence at the AND gates completing the requirement for coincidence logic.

The trip paths control the operation of trip curcuit breakers in the trip function block. When the measured pressure level is within its prescribed operating range, the trip circuit breakers are closed connecting the electric power supplied by the motor-generator sets to the control rod drive. If a protective action is initiated, then the trip paths will cause the circuit breakers to open interrupting electric power to the motive drive and initi-

ating protective action.

Other functional blocks indicated in the figure which interface with the system are power supply, motor-generator set, control rod motive drive, and TEST. They are included on the diagram to identify other systems which can be analyzed separately and evaluated collectively with the RPS.

Boundary of System and Level of Analysis

The boundary of the system for the analysis is shown by the dotted line in the functional block diagram (Fig. 3). All blocks contained within the dotted line were included in the study and those outside were excluded. "Those outside" means external interfacing items rather than components of the system itself.

The level of detail of the analysis is to the replaceable component level in the system. Some of these components can be represented by a piece part while others are represented by functional black boxes. An example of these parts is shown in Table I. The sensor units and trip function, for example, are analyzed as a functional "black-box". The development of their failures modes to the next lower level of component detail is not necessary to adequately assess system performance.

The failure mode of the component within each block was examined for its effect on the block function and then assessed for the overall system. The overall system can be evaluated by examining the performance of each block and then collectively looking at the overall system. The description of these blocks and their primary function are given in Table II.

Results

The results of the analysis provided useful feedback to the designer on the capabilities of this system. From the FMEA and Table of Failure Mode probabilities, a list of components whose most likely and least likely failure modes cause system effects was developed. This information gives the designer guidance in selecting circuits for preferred failure modes and in evaluating alternate design approaches. The means by which failures are detected or annunciated were indicated to the designer, and where necessary, additional surveillance procedures were added to detect unannunciated failures.

Conclusion

The failure modes and effects analysis is a useful technique for testing the protection system analytically. This method gave us information about the design to assess our compliance to the criteria. From this information, we learned how the system would work if failures occurred and how these failures can be detected. The results showed the expected performance of our system for degraded conditions. It identified which parts could cause problems so that design changes could be effectively made. The FMEA provided a means for doing an independent design review of our protection system as well as to document the results.

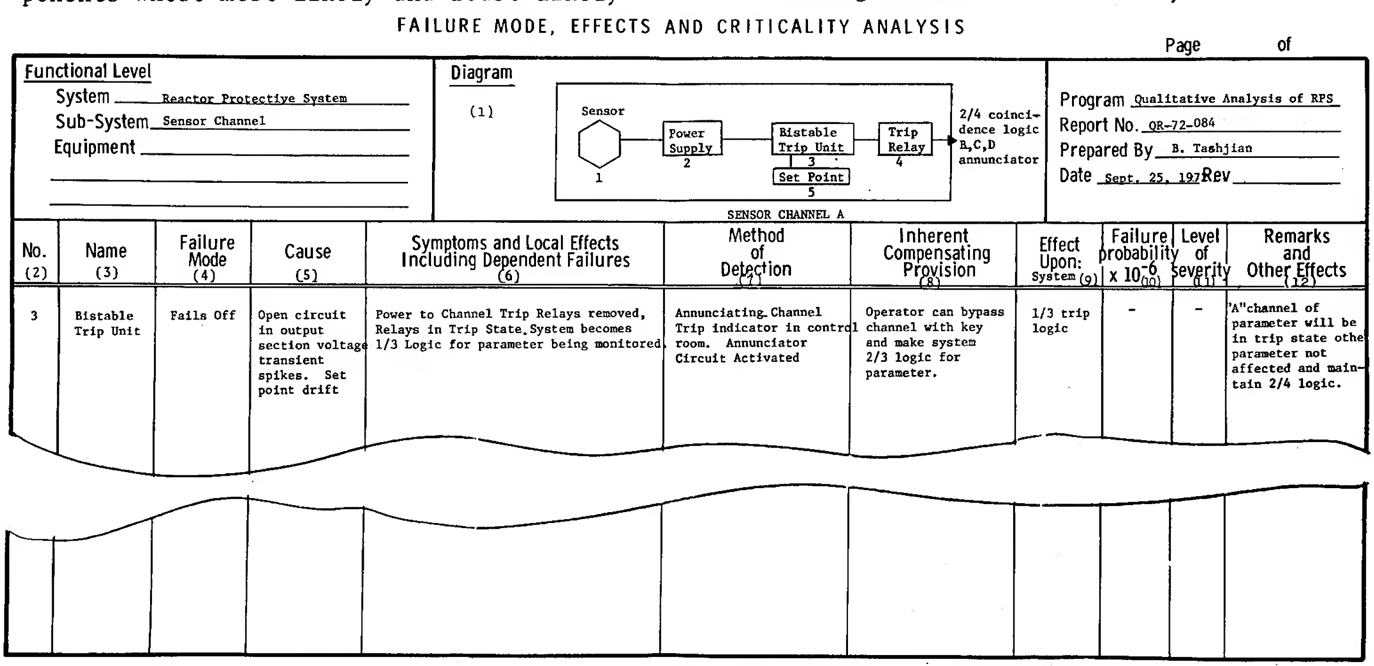
References

1. "General Design Criteria for Nuclear Power Plants, 10CFR50, Appendix A", February 20, 1971

2. ANSI N42.7/IEEE-STD-279, Criteria for Protection Systems for Nuclear Power Generating Stations (1971)

3. ANSI N41.4/IEEE-STD-352, IEEE Trial-Use Guide, General Principles for Reliability Analysis of Nuclear Power Generating Station Protection Systems

4. JCNPS/SC 1.1, Joint Committee on Nuclear Power Standards, Institute of Electrical and Electronic Engineers, IEEE Trial-Use Guide for the Application of the Single-Failure Criterion to Nuclear Power Generating Station Protection Systems



FAILURE MODE, EFFECTS AND CRITICALITY ANALYSIS

Page of

Functional Level
System Reactor Protective System
Sub-System Sensor Channel
Equipment

Diagram (1)

Program Qualitative Analysis of RPS
Report No. QR-72-084
Prepared By B. Tashjian
Date Sept. 25, 197 Rev

No. (2)	Name (3)	Failure Mode (4)	Cause (5)	Symptoms and Local Effects Including Dependent Failures (6)	Method of Detection (7)	Inherent Compensating Provision (8)	Effect Upon: System (9)	Failure Probability x 10^{-6} (10)	Level of Severity (11)	Remarks and Other Effects (12)
3	Bistable Trip Unit	Fails Off	Open circuit in output section voltage transient spikes. Set point drift	Power to Channel Trip Relays removed, Relays in Trip State. System becomes 1/3 Logic for parameter being monitored	Annunciating Channel Trip indicator in control room. Annunciator Circuit Activated	Operator can bypass channel with key and make system 2/3 logic for parameter.	1/3 trip logic	-	-	"A" channel of parameter will be in trip state other parameter not affected and maintain 2/4 logic.

Fig. 1: Failure modes and effect analysis worksheet

Table I

Generic Part Failure Rate and Failure Mode

Component Identification	Failure Rate f/hr	Open	Short	Failure Mode High	Low	Off	On	Other
AC Control Relay	5.0	68% (trip)	32% (trip)					
DC Power Supply	11.09			20%	20%	60%		
Circuit Breaker	5.0	42%	35%	(unable to reset 23%)				

Part	Type	Military Equivalent	Failure Rate $x10^{-6}$ (GF)*	Failure Mode** Open	Short	Other
Resistor	Fixed, Film	MIL-R-10509	.026	.90	.10	DRIFT
Resistor	Fixed, Composition	MIL-R-11	.033	.90	.10	
Resistor	Variable, CERMET	MIL-R-22097	.100	.95	.05	
Diode	Silicon, Diffused Junction	MIL-S-19500	.290	.30	.70	
TSTR	NPN, Epitaxial, Silicon	MIL-S-19500	.480	.65	.35	
TSTR	PNP, Expitaxial, Silicon	MIL-S-19500	.480	.65	.35	

*GF factor based on following failure rate source of military electronic component: MIL-HDBK-217A; UKAEC, AHS R(S) R117; 1968 Reliability Symposium; ARINC; Boeing Co; Battelle

**Failure modes based on following sources: NASA; FARADA; Battelle; UKAEC, AHSB R(S) R-117

Table II

Name	Function
System	Initiate trip function when pressure limit is exceeded
Trip Function	Interrupt power from motor-generator to rod control power supply
Trip Path	Break circuit to trip breaker UV coil on trip
2/4 Coincidence Logic	Breaks circuit to DC relays in trip path for any 2 of 4 inputs
Sensor Channel	Initiate trip signal for channel
Alarm Unit (Bistable, set point, trip relay)	Remove AC power to relays for $P_m > P_{set}$
Power Supply	Provide power for analog current loop
Sensor (Pressure transmitter)	Convert pressure to analog current

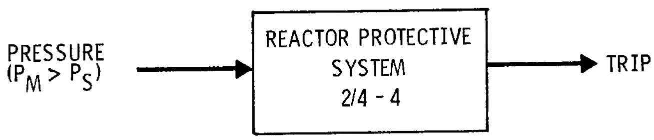


Fig. 2: Simplified protection system

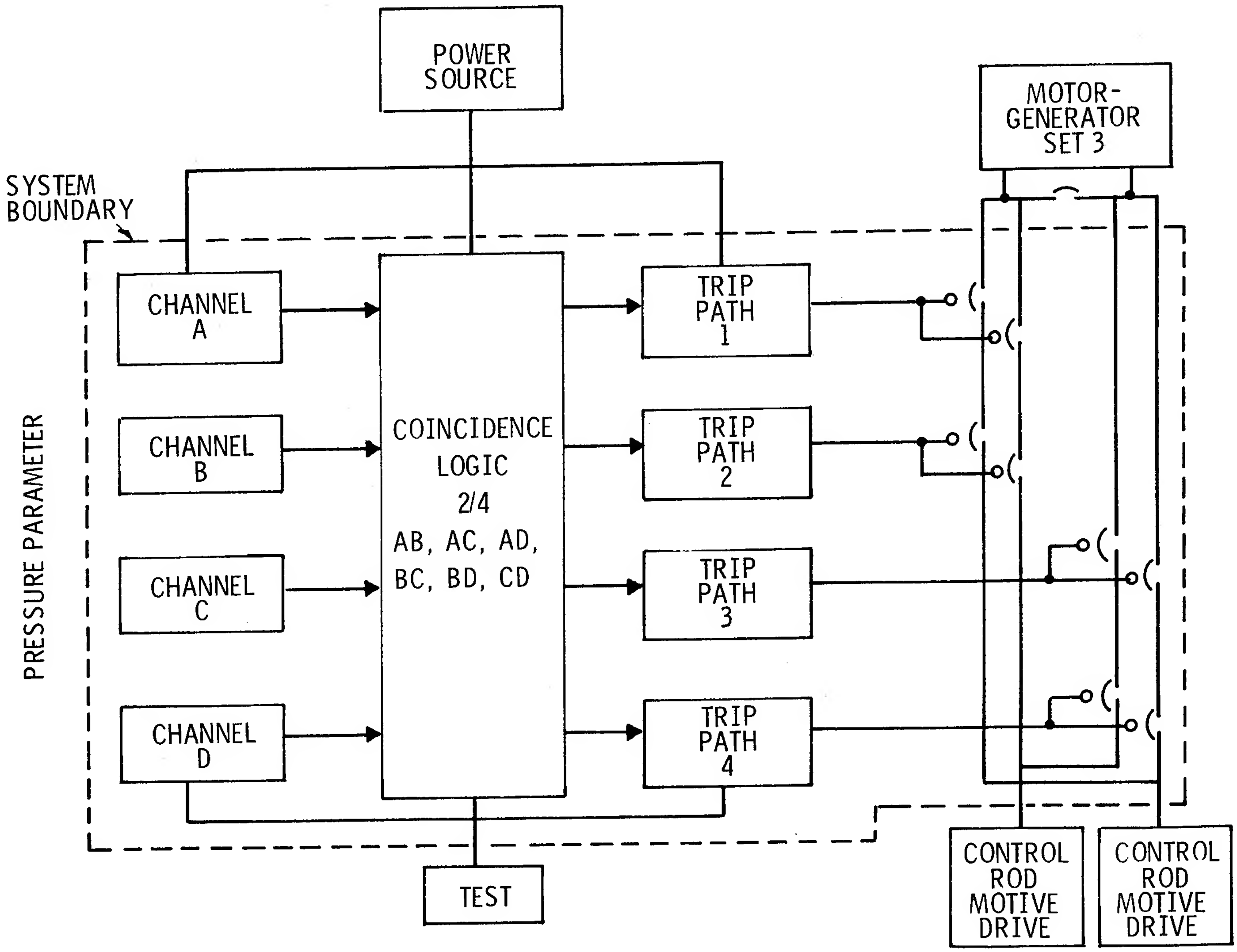


Fig. 3: Simplified functional diagram

EDISON ELECTRIC INSTITUTE NUCLEAR DATA COLLECTION SYSTEM

J. P. Whooley
Public Service Electric and Gas Company
Newark, New Jersey

Summary

The Equipment Availability Task Force of Edison Electric Institute (EEI) is in the final stages of establishing a reliability data collection system to analyze failure data from nuclear power plants.

The purpose of this paper is to describe the programs and procedures being established by EEI to provide the nuclear power industry with failure data statistics on nuclear plant components.

Introduction

Power plant reliability data has been collected, processed, and analyzed by the EEI Prime Movers Committee since the 1930's - first manually, and later by computer. An increasing demand for more detailed outage data in nuclear plants led the Prime Movers Committee's Equipment Availability Task Force to propose to the EEI Board of Directors that a computer system be established to collect component reliability data for the nuclear power industry.

This research project, designated as RP-101, was begun in 1971 and is now undergoing field test at a nuclear power plant in the mid-west. It is expected that the system will become fully operational during the first half of 1973.

EEI will be the focal point for the program which will involve the reactor manufacturers, electric utilities and other interested parties. At the present, there are approximately 30 utilities with nuclear plants operating or on order. By 1980 there will be 100 units in operation, using reactors supplied by five major companies currently manufacturing nuclear steam supply systems (NSSS).

The collection of reliability data will be confined to components of the reactor safety system, reactor protection systems and safety related systems. Reliability data will be available on both a component and a system basis. The data collection system will be operated by Southwest Research Institute under the direction of a special Steering Committee from EEI's Equipment Availability Task Force.

Scope

Data will be collected from any reactor system which is used primarily for the purpose of generating electric power. Failure data will be accepted from investor-owned utilities, cooperatives, municipal utility districts, power authorities, and the Atomic Energy Commission, as long as the facility is operated for the purpose of generating electricity. It is intended that the processed failure statistics and data on generic components would be available to the general public. For the present, the collection of data will be confined to organizations in the United States of America.

There are 3000 to 3500 components per generating unit for which failure reports must be submitted. Of these, there may be only 600 significantly different items. That is, there might be 600 pedigreed items with an average population of six per pedigree. It is estimated that there will be 50 failure reports per year generated for each unit.

Data Collection

In general terms, failure data will be reported on the components of the protection or safety systems which are installed to prevent or mitigate the consequences of a nuclear incident in the reactor system, but not on the structural components such as reactor vessels, containments, piping systems, buildings, supports, or mounting hardware, or on certain electronic or electrical parts such as fuses, resistors, diodes, transistors, etc.

Table P-1 summarizes the nuclear protection and safety systems in PWR and BWR nuclear steam supply systems. Table P-2 is a list of the components in these protection systems for which failure data will be reported.

The companies will report three types of data for each nuclear generating unit:

1. Pedigree Reports

   Pedigree (design) data will be submitted for each discrete component and/or system. This data is only submitted once, at the time the unit goes into service. (The computer input form plus a brief explanation of each field can be found in Appendix-P.)

2. Failure Reports

   Failure data on components and/or systems will be submitted on a quarterly reporting schedule. (The computer input form plus a brief explanation of each field can be found in Appendix-F.)

3. Quarterly Reports

   Plant operating information is reported on a single report form at the end of each quarter along with the Failure Reports. This report is used to update

the service hours of the system and components contained in the data base. (The computer input form plus a brief explanation of each field can be found in Appendix-Q.)

Data Base

The file maintenance system will run quarterly. Routines are provided to add, replace, correct, expand, or delete any record in the main file, including reports of failures.

New data will be validated before entry into the main file. Invalid information shall not be entered. Each utility shall correct any errors arising from submitted data for which it is responsible. Forms that are improperly filled out shall be returned to the point of origin with cause of return delineated.

Initially, and at any time when revisions are made, a complete listing of pedigree data for all specified components shall be made and sent back to the reporting company by EEI. Input source documents and the error listings will be mailed to the reporting utilities and the manufacturers after the quarterly file maintenance run.

Annual Reports

1. Protection System Reliability Report

An annual report will be published by EEI giving the reliability statistics for both the latest year and cumulative. The Report will list the reliability statistics shown below, for each type of reactor, summarized by system.

Unit Operating Hours

Plant Operating Hrs:
Plant Standby Hrs:
Plant Outage Hrs:

System Availability Data *

System I.D. (By EEI Designation)
System Mode of Functional Availability
Unit - Total Units Reporting
Population - Total Plants Same System
Calendar Hrs. Reporting Period/System
Total System Hrs. Available/System (Hrs)
Avg. System Availability/System (%)
Avg. System Outage Hrs/System (Hrs)
Avg. Duration Between Failures)
**Failure Rate/Population
**Failure Rate/Availability (Hrs)
No. of Failure Reports
Total System Hrs Outage/System (Hrs)

*Data presented per period, and cumulative
**Special Reports Only

2. Nuclear Unit Reliability Report

Each reporting organization will be given an annual report summarizing the availability of the protection systems for each nuclear unit reported to EEI. This report will list the component failures and associated data, including the effect, mode, type and cause of failure as shown below.

Utility Plant I.D.

Plant Operating Hrs:
Plant Standby Hrs:
Plant Outage Hrs:

System Availability Data*

System I.D. (By EEI Designation)
System Mode of Functional Availability
System I.D.
Population
Calendar Hrs Reporting Period
Total Hrs Available/System (Hrs)
No. of Failure Reports
Total Hrs Outage/System (Hrs)
Avg. System Availability/System (Hrs)
Avg. System Outage Hrs/System (Hrs)
(Avg. Duration Between Failures)
** Failure Rate/Population
** Failure Rate/Availability Hrs

Systems/Components Failure Data Listing

System/Components I.D.
Component I.D. by Utility Designation
Date of Failure
Failure Outage Duration (Hrs)
Applicable Protective System I.D.
Failure Condition/Action Codes

3. Special Reports

Properly authorized special requests for analysis of the data base will be honored. EEI will supply the request forms plus instructions. Information will be retrievable on any combination of pedigree fields, event fields, or control fields and on selected portions of these fields. All pedigree data shall be usable for retrieval and interrogation with a combination of sorts, certain range selection and logical combinations. Failure rates, average availability, average outage duration, etc. may be requested on these Special Reports. In addition, each utility and NSSS vendor may request a complete copy of all his pedigree and failure data on tape or as a listing.

Conclusion

The nuclear utility industry has been provided with a valuable tool for analyzing and improving the reliability of safety and protection systems for nuclear steam supply systems. With the tool comes the challenge to use it well. This can only be achieved if all segments of the industry give the new data collection system their full cooperation.

P

EEI NTSC 7203 FORM EEI-3

NUCLEAR PLANT RELIABILITY DATA SYSTEM ;

REPORT OF PEDIGREE

This data entry is for:

☐ New Pedigree Data ☐ Replacement Pedigree

☐ Correction to Pedigree

INPUT CONTROL	CODE	UTILITY DESIG.	PLANT	UNIT	EEI COMP. NAME (OR) $ $ $ SYSTEM	UTILITY PLANT EQUIPMENT IDENTIFICATION NUMBER BASE	COMPONENT I.D. NO:	SUB NO	PEDIGREE DATE (FLOAT)		
	1 2 3	4 5 6	7 8 9	10	11 12 13 14 15 16	17 18	19 20 21 22 23 24 25 26 27 28	29 30	31 32 33 34 35 36 (YR MO DY)	37	INPUT CONTROL DATA
	C 7 E									P	

SYSTEM / COMPONENT ENGINEERING DATA	38–80			
A	SYSTEM / COMP. TITLE /	71	UTL-SYSTEM-CLS.	EEI-SYST.
B	QUAN. -UNIT 45 PART IDENTIFICATION NO / VENDOR EQUIP. NO S/M 64 SERVICE DATE Y M D	71	MODE ★ 75	ENVIRONMENT
C	A 41 B 44 C 47 D 50 E 53 F 56 G/SIZE 62 H/ -UNITS	71	J/ -UNITS	
D	SOURCE: (SUPPLIER / VENDOR 56 MANUFACTURE:	71		
E				

PEDIGREE UPDATE	38–59				
	OPTIONAL % USE IN MODE FUNCT.			TEST	
F	1-REACTOR WHILE CRITICAL >2% POWER	2- STANDBY CONDITION	3- REACTOR OPERATIONS SHUTDOWN	4-TECH. SPEC. TESTS PER QUARTER (PEDIGREE)	59
	. %	. %	. %		

APPROVE	G	INIT:	DATE Y M D	ACTION APPROVAL: SIGN	59

★ MODE:

FDP	-	Functional During >2% Power
SDP	-	Standby During Power
SDC	-	Shutdown Condition During Power

ENVIRONMENT:

APPENDIX P

REPORT OF PEDIGREE
PROCEDURE

A. Form Entry Instructions

FIELD	LINE/COL.	INSTRUCTION
Utility Desig.	Control (4-6)	Enter EEI designated Utility code
Plant/Unit	Control (7-10)	Enter EEI designation of Plant code
Comp/ Systems Name	Control (11-16)	Enter EEI designated system code or component name for Tables P-1, and P-2, respectively.
Base (Comp-only)	Control (17-18)	As above from Tables P-2 designating base files category for components only.
Equipment Number (Comp-only)	Control (19-30)	Enter Utilities plant equipment number as used by the Utility to identify the specific component.
Pedigree Date (FLOAT)	Control (31-36)	a. Pedigree Date, is flagged by the word (FLOAT) which is a necessary warning that the Pedigree Date established on this report has a "floating" capability that dictates the programs computational routines by establishing the statistical start date for calculations against that system or component. b. Enter that date in time on which the system or component data reliability calculations will begin accounting. c. This date should be based not on service date but that date where all accountable failure history records provide liquidate accounting for available service.

Syst/Comp Name	A (39-70)	Enter the Utility's System or Component name.
Utl-System	A (72-75)	Enter the Utility's System Code designation (starting from left hand side of field).
CLS	(A) (76-77)	
EEI Syst	(A) (78-80)	Enter EEI designated System Code.
Quan	(B) (39-42)	Enter quantity or number of units represented by the "Utility Plant Equipment I. D. No." Generally, unit is only one (1), however, where multiple units exist, and Utility Equipment I. D. No. is not sub-numbered to give individual (unique) designation to each component, use quantity of units represented. Systems pedigree should be uniquely identified by separate system code even if a redundant system.
Unit	(B) (43-44)	Unit should be entered as (EA-for each), (FT-for feet) etc., as applicable.
Part I. D.	(B) (46-62)	Enter part identification no. using designation selected to insure a common part identity that will allow the Utility to reorder and identical part, regardless of the original source, so that source can trace it to the manufacturer and manufacturers specification and/or fabrication drawing.

Part I. D. S/M	(B) (63)	Enter S or M designating the part identification no. origin as that of the S . Supplier or M - manufactures listed in line D.
Service Date	(B) (65-70)	Enter actual service date of System or component date sequence is by YR-MO-DAY.
Mode	(B) (72-74)	Enter system or component mode of operation which exists during normal operating conditions above reactor is critical over 2% power. See Mode - Codes on Pedigree form.
Environment	(B) (76-80)	

ENGINEERING DATA

Engr. Data	(C) (39-80)	Enter applicable engineering data, codes and values listed on Table P-3
Source	(D) (39-55)	Enter name of the source of the system or component identifying the supplier or vendor.
Manuf.	(D) (57-70)	Enter the name of the systems or component manufacturer even if manufacturer is the same as supplier.
Pedigree Update	(F-1, 2, 3) (39-53)	Enter (Optimal) % of time the system or component will be operating in the specified MODE function 1, 2, or 3, as an expected normal function condition during reactor operation over 2% power.
Pedigree Update (test)	(F-4) (54-58)	Enter the number of scheduled tests (Periodic Testing) per quarter for that system or component as defined in the plant's Technical Specifications.
Approve	(G)	Enter individuals' initials, date of approval, and signature of one who certifies approval of a submitted report.

APPENDIX P

Q

EEI NUCLEAR PLANT RELIABILITY DATA SYSTEM ;

NTSC 7203
FORM EEI-4

QUARTERLY REPORT

INPUT CONTROL	CODE	UTILITY DESIG.	PLANT	UNIT		REPORT DATE		INPUT CONTROL DATA
	1 2 3	4 5 6	7 8 9	10	11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	31 32 33 34 35 36 (YR MO DY)	37	
	C 7 E				$ $ $ $ $ $		P	

QUARTERLY REPORT UPDATE	A	PLANT NAME							
	38	39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58				59	60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80			
		QUARTERLY UPDATE DATA					QUARTERLY REPORT VERIFICATION DATA			
	F	1-REACTOR HRS. WHILE CRITICAL >2% POWER	2-STANDBY CONDITION (HRS)	3-REACTOR OPERATIONS SHUTDOWN (HRS)		59	QUARTER START DATE: Y M D	QUARTER END DATE: Y M D	X-THIS QUARTER 1 2 3 4	FAILURE RPTS. THIS QUART. SYST 78 COMP
APPROVE	G	INITIAL	DATE Y M D	ACTION APPROVAL: SIGN						

APPENDIX Q

QUARTERLY REPORT PROCEDURE

A. Form Entry Instructions

FIELD	LINE/COL.	INSTRUCTION
Utility Desig.	Control (4-6)	Enter EEI designated Utility code
Plant/Unit	Control (7-10)	Enter EEI designation of Plant code
Report Date	Control (31-36)	Enter date report is made, preferably last date of designated quarter. Date sequence is by YR-MO-DAY.
Plant Name	A (39-80)	Enter the Utilities' Plant or Unit name.
Update Data	(F-1) (39-43)	a. Enter number of calendar hrs reactor was critical over 2% power for that quarter. b. The first Quarterly Report submitted should reflect total accumulative hrs to date to account for prior history reactor critical hrs over 2% power.
Update Data	(F-2) (44-48)	a. Enter number of calendar hrs reactor was in standby condition during that period. b. The first Quarterly Report submitted should reflect total accumulative hrs to date to account for prior history reactor standby condition.

FIELD	LINE / COL.	INSTRUCTION
		d. This date should be (service date) for any newly installed component.
Update Data	(F-3) (49-53)	a. Enter number of calendar hrs reactor was in a shutdown condition during that quarter. b. The first Quarterly Report submitted should reflect total accumulative hrs to date to account for prior history. c. Note: in completing positions F 1, 2, and 3, assure that the total hrs of F 1, 2, and 3 equal to the actual calendar hrs in that quarter.
Quarter Start Date	(F) (60-65)	Enter date of the first day of the designated quarter. Date sequence is by YR-MO-DAY.
Quarter End Date	(F) (66-71)	Enter date of the last day of the designated quarter. Date sequence is by YR-MO-DAY.
This Quarter	(F) (72-75)	Enter "X" in the designated calendar quarter field.
RPTS - This Quarter	(F) (79-80)	Enter: a. Number of Systems Failure Reports submitted in report package to EEI this quarter. b. Number of Components Failure Reports submitted in report package to EEI this quarter.
Approve	(G) (38-39)	Enter individuals' initials, date of approval, and signature of one who certifies approval of submitted report.

F

EEI
NTSC 7203
FORM EEI-3

NUCLEAR PLANT RELIABILITY DATA SYSTEM ;

REPORT OF FAILURE

This data entry is for:

☐ - New failure report data.

☐ - Correction to previous submitted failure data with same control identification shown.

KEYPUNCH NOTE

1. ITEMS 1 THRU 37 ON FRONT SIDE DESIGNATES INPUT CONTROL DATA, AND MUST BE KEY-PUNCHED BEFORE ALL ITEMS LISTED BELOW.
2. [ ] AND ▒ = KEYSKIP (BLANKS)

INPUT CONTROL	CODE	UTILITY DESIG.	PLANT	UNIT	EEI COMP. NAME (OR) $ $ $ SYSTEM	UTILITY PLANT EQUIPMENT IDENTIFICATION NUMBER: BASE	COMPONENT I.D. NO:	SUB NO	ACTION ENTRY DATE YR MO DY	S OR C	S-SYSTEM FAIL / C-COMP. FAIL
	1 2 3	4 5 6	7 8 9	10	11 12 13 14 15 16	17 18	19 20 21 22 23 24 25 26 27 28	29 30	31 32 33 34 35 36	37	INPUT CONTROL DATA
	C 7 E										

EVENT DATA	38	39–48	49	50–59	60	61–66	67	68–70	71	72–77	78–80
	E	FAILURE EVENT START Y M D HR MIN	49	FAILURE EVENT END Y M D HR MIN	60	UTL. SYSTEM-CLS:	67	EEI-SYST.	71	EEI COMP. NAME	

FAILURE DESCRIPTION (75-CHAR.)	H	
	J	
CORRECTIVE ACTION (75-CHAR.)	K	
	L	

EEI FAILURE DATA

1. CIRCLE COLUMN CODES UNDER EACH APPLICABLE FAILURE CLASSIFICATION
2. CIRCLE AS MANY COLUMN CODE DESIGNATIONS AS NECESSARY TO ADEQUATELY DESCRIBE FAILURE SITUATION

	MODE OF FAILURE						TYPES OF FAILURE							CAUSE OF FAILURE								
	WON'T START /CLOSE	WON'T STOP /OPEN	WON'T OPER. PER. SPEC.	SPURIOUS OPERATION	FALSE RESPONSE		ELECTRICAL	MECHANICAL	METALLURGY	CORROSION	INSTRUMENT.	OTHER		ENGR/DESIGN	MANUFACT.	CONSTR./ INSTALL	OPERATING ERROR	MAINT./ TESTING	UNUSUAL SERV. COND.	NATURAL END LIFE	OTHER	UNKNOWN
38	39 40	41 42	43 44	45 46	47 48	49	50 51	52 53	54 55	56 57	58 59	60 61	62	63 64	65 66	67 68	69 70	71 72	73 74	75 76	77 78	79 80
V	MA	MB	MC	MD	ME		TA	TB	TC	TD	TE	TF		CA	CB	CC	CD	CE	CF	CG	CH	CI

	FAILURE DETECTION						ACTION TAKEN							EFFECT OF FAILURE								
	TESTING	PREVENT. MAINT.	AUDIO/VISUAL ALARM	ROUTINE INSPECT.	OTHER		REPLACE TOTAL COMP.	REPLACE PART/S	REPAIR PART/S	RECALIB/ ADJUST.	REDESIGN MODIFY	NONE		LOSS OF SYST FUNCTION	SYSTEM OPER. DEGRADED	PREVENTS COMP. OPER.	LOAD REDUCTION	DAMAGE TO OTHER COMP.	SCRAM	OTHER SYST. DEGRADED	GEN. UNIT OFF SYSTEM	NO EFFECT
38	39 40	41 42	43 44	45 46	47 48	49	50 51	52 53	54 55	56 57	58 59	60 61	62	63 64	65 66	67 68	69 70	71 72	73 74	75 76	77 78	79 80
W	DA	DB	DC	DD	DE		AA	AB	AC	AD	AE	AF		EA	EB	EC	ED	EE	EF	EG	EH	EK

APPROVE	G	INIT:	DATE Y M D	ACTION APPROVAL: SIGN	60	INIT:	DATE Y M D	Q.A. AUDIT: SIGN

REPORT OF FAILURE
PROCEDURE

A. Form Entry Instructions

FIELD	LINE/COL.	INSTRUCTION
Utility Desig.	Control (4-6)	Enter EEI designated Utility Code
Plant	Control (7-10) (7-10)	Enter EEI designation of Plant Code
Comp/ Systems Name	Control (11-16)	Enter EEI designated system code or component name for Tables P-1, and P-2, respectively.
Base (Comp-only)	Control (17-18)	As above from Tables P-2 designating base files category for components only.
Equipment Number (Comp-only)	Control (19-30)	Enter Utilities plant equipment number as used by the Utility to identify the specific component.
Action Date	Control (31-36)	a. Enter the date on which the system or component failure occurred. Date sequence is by YR-MO-DAY.
S/C	Control (37)	Enter S - if Systems Failure Report C - if Component Failure Report

Failure Start	(E) (39-48)	a. Enter date and time failure event started including exact or approximate time. Sequence is by YR-MO-DAY-TIME. b. Time may be required to designate date failure event was detected, if actual event start is not known.
Failure Start	(E) (50-59)	Enter date and time failure event ended as of the time the system or component was placed back into available service.
Utl-System	(E) (61-66)	Enter the Utility's System Code designation (starting from the left hand side of field).
CLS	(E) (65-66)	
EEI Syst	(E) (68-70)	Enter EEI designated System Code.
EEI Comp.	(E) (72-77)	Enter EEI - Component name from Table P-2, which caused the failure. This applies to both system and component failure.
Failure Descript	(H & J) (39-80)	Indicate by large character printed statement and description of the failure.
Correct. Action	(K & L)	Indicate by large character printed statement and description of the corrective action taken.
EEI Failure Data	(V & W) (39-80)	1. Circle column codes under each applicable failure classification. 2. Circle as many column codes designations as necessary to adequately describe failure situation.
Approve	(G)	Enter individuals' initials, date of approval, and signature of one who certifies approval of a submitted report.

TABLE P-1 - PWR
EEI NUCLEAR SYSTEMS CLASSIFICATION

CODE	SYSTEM TITLE	CODE	SYSTEM TITLE
RVG	Reactor Vessel General	PSR	Primary System Relief
RVI	Reactor Vessel Internals	WDS	Waste Disposal System
RCS	Reactor Coolant System	RHR	Residual Heat Removal System
RPS	Reactor Protection System	CCS	Component Cooling System
ACS	Auxiliary Coolant System	CVC	Chemical and Volume Control
ECC	Emergency Power System	ESF	Engineered Safety Features
EPS	Emergency Power System	FHS	Fuel Handling System
ROD	Control Rod System	FHC	Fuel Handling Crane
BIS	Boron Injection System	ICI	Incore Instrumentation
CIS	Containment Isolation System	NPC	Nuclear Process Control and Instrumentation
CSS	Containment Spray System	IVS	Isolation Valve Seal Water System
CAR	Containment Air Removal System	CWS	Criculating Water System
CLP	Containment Liner Penetration	RMS	External Rad Monitoring System
CRS	Containment Recirculation System	RWS	Refueling Water Storage
SIS	Safety Injection System	SAM	Sampling System
PRE	Pressurizer	SFC	Spent Fuel Pit Cooling
PRS	Pressurizer Relief System	EMF	Emergency Boiler Feed
AAS	Associated Auxiliary Systems		
GEN	Steam Generator and Associated Systems		

TABLE P-1 - BWR
EEI NUCLEAR SYSTEMS CLASSIFICATION

CODE	SYSTEM TITLE	CODE	SYSTEM TITLE
RVG	Reactor Vessel General	ADS	Auto-Depressurization System
RVI	Reactor Vessel Internals	WDS	Waste Disposal System
RCS	Reactor Coolant System	LPC	Coolant Injection System (Lo-Press)
RPS	Reactor Protective System	RHR	Residual Heat Removal System
ACS	Auxiliary Coolant System	CCS	Component Cooling System
ECC	Emergency Core Cooling	CVC	Chemical and Volume Control
EPS	Emergency Power System	ESF	Engineered Safety Features
ROD	Control Rod System	FHS	Fuel Handling System
CIS	Containment Isolation System	FHC	Fuel Handling Crane
CSS	Containment Spray System	ICI	Incore Instrumentation
CIR	Containment Inerting System	CTR	Control Equipment for Class I Equipment
CLP	Containment Liner Penetration	NPC	Nuclear Process Control and Instrumentation
CVS	Containment Ventillation System	IVS	Isolation Valve Seal Water System
SIS	Safety Injection System	CWS	Circulating Water System
LCS	Liquid Control System (Standby)	RMS	External Rad Monitoring System
GTS	Gas Treatment System (Standby)	RWS	Refueling Water Storage
AAS	Associated Auxiliary Systems	SAM	Sampling System
HPC	Coolant Injection System (Hi-Press)	SFC	Spent Fuel Pit Cooling
PSR	Primary System Relief	EMF	Emergency Boiler Feed, Service Water, and Fire Protection Systems' Pumps and Piping

TABLE P-2

LIST OF COMPONENTS

Description	Comp. Name	File Base
Amplifiers	AMPLIF	IN
Annunciators	ANNUNC	IN
Batteries	BATTRY	EE
Circuit Breakers	CIRBRK	EE
Contactors, Starters	CONTAC	EE
Demineralizer	DEMINR	ME
Engines, Internal Combustion	ENGINE	ME
Fans/Ventillators/Coolers	FANVEN	ME
Filter/Strainers	FILTER	ME
Generators	GENERA	EE
Heat Exchangers	HTEXCH	ME
Modules/Elements	MODELM	EE
Motor, Electric	MOTORS	EE
Power Supplies	POWSUP	EE
Preamplifiers	PREAMP	IN
Pressurizers	PRESUR	ME
Pumps	PUMPGN	ME
Radiation Monitors	RADMON	ME
Regulators	REGULA	EE
Relays	RELAYS	EE
Sensor, Flow	SENFLO	IN
Sensor, Level	SENLEV	IN
Sensor, Pressure	SENPRE	IN
Sensor, Temperature	SENTEM	IN
Steam Turbines	TURBIN	ME
Switches	SWITCH	EE
Transformers	TRANSF	EE
Valves	VALVES	VL

DESIGN CRITERIA—WHICH ONE?

Clifford B. Boehmer
McDonnell Douglas Astronautics Company
Huntington Beach, California

Introduction

The subject of design criteria has come up on several recent space programs, and specific criteria have been imposed on some of them. An interesting and vital question is usually raised when this happens—"which criterion should be used for our project, and why?" The answer can be obtained by "gut feeling" techniques such as "we want a system that is still operable after the first failure," or "our last successful project had a reliability of 0.95 so why change," etc. But perhaps a better technique is to perform an economic tradeoff between the available criteria to determine the most cost effective. The results of this trade study can then be modified, if desired, to account for items that cannot be measured in terms of monetary values, such as the safety of astronauts.

The specific project used in the trade study described in this paper was a Reusable Nuclear Shuttle (RNS) with a 75,000-lb-thrust nuclear engine. This engine concept was replaced with a 15,600-lb-thrust nuclear engine early this year; however, the results of this study can be generalized to apply to a number of projects, including the new RNS with the smaller nuclear engine.

The RNS was a nuclear-powered space transportation system employing the Nuclear Engine for Rocket Vehicle Application (NERVA), which could be based in low (260 nmi) earth orbit and used for transporting large payloads to lunar or geosynchronous orbits.[1,2,3] A typical lunar mission involves the transfer of 127,000 lb of cargo and men to a 60-nmi lunar polar orbit and the return of 20,000 lb to the low-earth-orbit home base. The geosynchronous shuttle mission involves the transfer of 117,000 lb to the geosynchronous orbit (19,325 nmi) and a return of 20,000 lb.

Two basic RNS concepts, each compatible with the assigned mission requirements, have been identified and studied: a 33-ft-diameter concept, designated Class 1 (Figure 1), and a multi-module concept comprising 15-ft-diameter elements and designated Class 3 (Figure 2). Both concepts employ three types of modules; a propellant module, a propulsion module, and a command and control module. The propellant module(s) contains the bulk of the liquid hydrogen propellant, the propulsion module contains the NERVA engine and a small amount of liquid hydrogen propellant, and the command and control module contains the majority of the astrionics system, the electrical power system, and the auxiliary propulsion (attitude control) system. The Class 1 concept contains one propellant module plus one propulsion and one command and control module. The Class 3 concept contains eight propellant modules, one propulsion module, and one command and control module. Both concepts are assembled in earth orbit and have a usable LH_2 capacity of about 300,000 lb. The Class 1 propellant module is placed in earth orbit by a modified Saturn V launch vehicle. The Class 3 propellant modules and the common propulsion and command and control modules are placed in earth orbit by the space shuttle.

The RNS systems (propulsion, astrionics, etc.) are fairly representative of state-of-the-art space

WD1903
29423

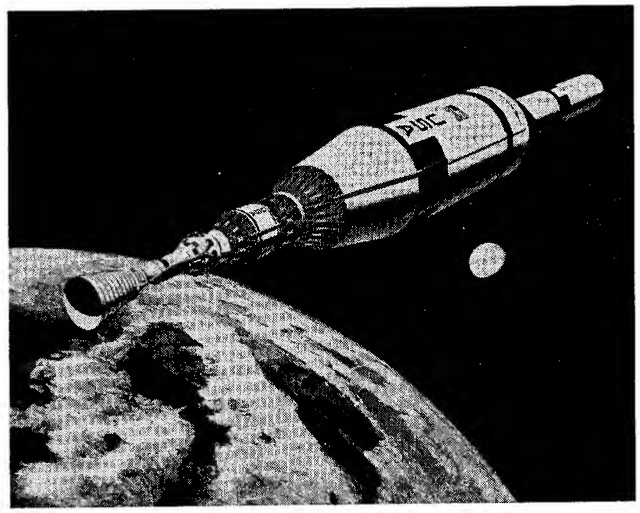


Figure 1. Nuclear Shuttle Hybrid-1

WD1903
29422

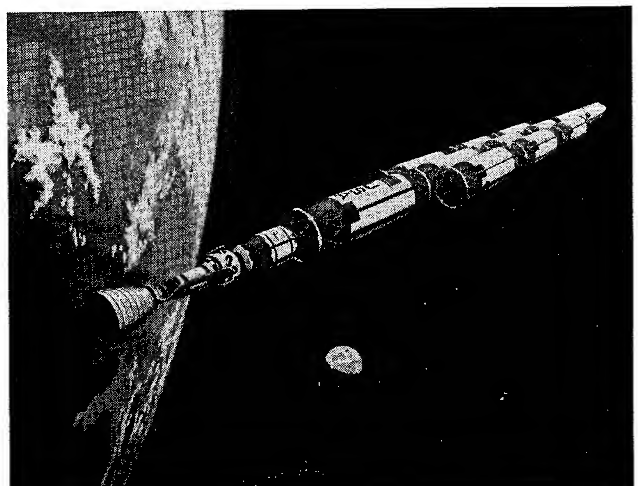


Figure 2. Class 3 Nuclear Shuttle

vehicles. The significant difference is the requirement for long life without maintenance for the propulsion and the propellant modules. The command and control module is returned to earth for maintenance and refurbishment after each of the ten round trips expected of a single RNS; the propellant module and the propulsion module must operate for the complete ten-mission tour without maintenance. If a module becomes inoperative it is discarded.

Procedure

The procedure employed in the trade study was to determine the benefit and associated cost of each of the following four general design criterion candidates:

1. Single—a single thread design with no redundancy.

2. No Single Failure (NSF)—a design where one failure will not cause a loss of the system.

3. Fail Operational-Fail Safe (FO-FS)—a design where the system is fully operational after the first component failure and is in a safe condition after the second component failure.

4. Fail Operational-Fail Operational-Fail Safe (FO-FO-FS)—a design where the system is fully operational after the first and second component failure and in a safe condition after the third component failure. This criterion was only applied to the astrionics systems; therefore, a hybrid criterion was actually used, FO-FS for all mechanical components and FO-FO-FS for all electronic components.

A direct comparison was achieved by calculating the overall cost of each candidate; i.e., the cost of meeting the criterion by adding components plus the cost of component failures. The former cost will increase as increasingly restrictive criteria are applied, but the latter cost will decrease. The optimum criterion for any specific application is defined as that for which overall cost is a minimum.

The first step in the study was to perform a multiple failure-mode analysis on the current RNS design (all non-structural components) to determine the effect of multiple failures. This analysis was then used to synthesize four systems that satisfied the four candidate criteria. The additional components for each of these synthesized designs was determined along with the reliability (mission success probability) and maintenance requirements. The cost (in dollars) for each of these three factors for a 60-mission tour (six RNS vehicles for ten missions each) was then determined. This included the cost of the added components, the cost of mission failures, and the cost of the maintenance operations. These costs were then added to obtain the overall cost for a specific design criterion.

System Analysis

A multiple failure mode effects analysis was performed on all non-structural components of the RNS systems.[4] Figure 3 shows a typical work sheet for this analysis. The analysis provided the data required to synthesize four systems that would satisfy (but not exceed) the four design criteria previously identified. Figure 4 shows a typical work sheet for the design synthesis task.

FAILURE MODE	MISSION PHASE	FAILURE EFFECT	FAILURE CLASS	REMARKS
ITEM 36.01.03.12 QUAD VALVES LH_2 TANK VENT ONE VALVE FAILS OPEN	F→V	NONE	FO	
	W	NONE	FO	RNS MUST BE REPAIRED
TWO SERIES VALVES FAIL OPEN	F→V	LH_2 IS LOST OVERBOARD. ATTITUDE CONTROL AND ELECTRIC POWER CAPABILITY IS LOST.	FU	
ONE VALVE FAILS CLOSED	F→V	NONE	FO	
	W	NONE	FO	RNS MUST BE REPAIRED
TWO PARALLEL VALVES FAIL CLOSED	F→V W	VENT CAPABILITY IS LOST. POSSIBLE TANK RUPTURE.	FC	

FAILURE CLASS
FO - FAIL OPERATIONAL
FS - FAIL SAFE
FU - FAIL UNSAFE
FC - FAIL CATASTROPHIC

Figure 3. Multiple Failure Mode Effects, System-APS 36.01.03

ITEM	NO SFI	FO-FS
CHECK VALVE 15.09.05	NONE	NONE
FLOW METER 15.09.03	NONE	NONE
SPRAY NOZZLE 15.09.06	NONE	NONE
GROUND FILL VALVE 15.09.07	1 VALVE IN SERIES (EXPLOSIVE CLOSED) OR A CHECK DISCONNECT	2 VALVES IN SERIES OR 1 VALVE IN SERIES (EXPLOSIVE) AND A CHECK DISCONNECT
GROUND VENT VALVE, 15.10.03 GROUND VENT AND RELIEF VALVE, 15.10.02 CHECK DISCONNECT, 15.10.05	NONE	1 VALVE IN SERIES WITH CHECK DISCONNECT (EXPLOSIVE CLOSED)
FLIGHT VENT QUAD VALVE, 15.11.02	NONE	A THIRD VALVE IN SERIES AND A THIRD VALVE IN PARALLEL (9 VALVES TOTAL)
CHILLDOWN PUMPS 25.05.01, 25.05.02 CHECK VALVES 25.05.03, 25.05.04	NONE	A THIRD SET (PUMP AND CHECK VALVE) IN PARALLEL OR ADD SUFFICIENT LH_2 TO ALLOW OPEN LOOP CHILL

Figure 4. Typical Design Synthesis

The four synthesized designs were then subjected to a multiple failure reliability analysis and a maintenance analysis.

The multiple failure reliability analysis determined the probability of a system failure causing loss of the mission in such a way that the RNS is unrecoverable, the payload lost, and the crew and/or passengers must be rescued.

The maintenance analysis used reliability techniques to determine the requirements for maintenance in earth orbit. The ground rules established for the RNS for this study and for the project were that the RNS cannot leave earth orbit with a single failure item (SFI), i.e., an item which, if it fails, could cause loss of the mission. Therefore, except for the single thread design, if the system contains a SFI in earth orbit, maintenance must be performed. The maintenance philosophy for the RNS, established by another trade study, is that there is no maintenance performed in earth orbit. After every trip, the command and control module is brought back to earth, where maintenance can be performed. A propulsion module or a propellant module containing a SFI is discarded, and a new module is brought up from earth.

The maintenance philosophy for the single thread design is that an RNS arriving in earth orbit with a failed component must be discarded. It must be pointed out that it is possible for a component to fail without loss of the mission, and this fact is recognized in the reliability and maintenance calculations.

The results of these analyses are shown in Tables 1 and 2 for the four types of modules—a Class 1 propellant module, a Class 3 propellant module, a propulsion module, and a command and control module. The mission failure probabilities and maintenance probabilities are expressed in terms of probability per mission. No penalty was charged to the command and control module maintenance because that maintenance is performed on the ground with no orbital operations required.

Economic Analysis

The cost involved in the addition of components to meet the more restrictive criteria is to a large extent reflected in the added weight. Tables 1 and 2 give the added weight imposed by each criterion. The cost of one pound of added weight can be obtained by calculating the reduction in payload or conversely the number of RNS vehicles that must be added to a 60-mission

Table 1
CLASS 1 RNS

MODULE	CRITERIA	TOTAL WEIGHT (LB)	DELTA WEIGHT (LB)	MISSION FAILURE PROBABILITY	MAINTENANCE PROBABILITY
PROPELLANT	SINGLE	29,755	0	0.003143	0
	NSF	30,206	451	0.000664	0.081680
	FO-FS	30,664	909	0.000610	0.005470
	FO-FS / FO-FO-FS	30,874	1,119	0.000608	0.004276
PROPULSION	SINGLE	33,195	0	0.042926	0.010325
	NSF	33,715	500	0.000669	0.136306
	FO-FS	34,266	1,071	0.000591	0.003348
	FO-FS / FO-FO-FS	34,580	1,385	0.000589	0.002151
COMMAND AND CONTROL	SINGLE	3,723	0	0.305474	—
	NSF	5,197	1,474	0.007943	—
	FO-FS	6,448	2,725	0.007515	—
	FO-FS / FO-FO-FS	7,219	3,496	0.007496	—

Table 2
CLASS 3 RNS

MODULE	CRITERIA	DELTA WEIGHT (LB)	TOTAL WEIGHT (LB)	MISSION FAILURE PROBABILITY	MAINTENANCE PROBABILITY
PROPELLANT	SINGLE	6,440	0	0.160674*	0 *
	NSF	6,883	443	0.004851	0.437794
	FO-FS	7,325	885	0.004569	0.028128
	FO-FS / FO-FO-FS	7,530	1,095	0.004559	0.022360
PROPULSION	SINGLE	30,795	0	0.042926	0.010325
	NSF	31,315	520	0.000699	0.136306
	FO-FS	31,866	1,071	0.000591	0.003348
	FO-FS / FO-FO-FS	32,180	1,385	0.000589	0.002151
COMMAND AND CONTROL	SINGLE	3,723	0	0.305474	—
	NSF	5,197	1,474	0.007943	—
	FO-FS	6,448	2,725	0.007515	—
	FO-FS / FO-FO-FS	7,219	3,496	0.007496	—

*8 MODULES

tour to achieve the same integrated payload to lunar orbit. The total payload to lunar orbit for the 60-mission tour using the Class 1 RNS is 7.62×10^6 lb at a total cost of $4,440.3 million or $582/lb. For every pound that is added to the RNS vehicle we lose 2.73 lb of payload per trip or 163.8 lb of payload for the 60 trips. Therefore that one pound of payload results in a total cost of $95,500. This value far outweighs any costs associated with the purchase of the additional components and additional engineering.

The cost of mission failures was determined by adding the cost of replacing the lost RNS, including the cost of the modules themselves ($54.5 million), the cost of launch vehicles ($96.1 million), launch operations ($57.7 million), ground support and engineering operations ($26.1 million), mission operations ($3.6 million), the value of the payload ($127 million), and the cost of crew rescue ($5.8 million). This results in a total cost of $370.8 million. This value is multiplied by the expected number of vehicles lost over the 60-mission tour as shown in Table 3.

Table 3
CLASS 1 RNS

MODULE	CRITERIA	WEIGHT EQUIVALENT VEHICLES	WEIGHT PENALTY ($10⁶)	MISSION FAILURE EQUIVALENT VEHICLES	MISSION FAILURE PENALTY ($10⁶)	MAINTENANCE EQUIVALENT MODULES	MAINTENANCE PENALTY ($10⁶)	OVERALL TOTAL COST ($10⁶)
PROPELLANT	SINGLE	0	0	1.886	699.1	0	0	699
	NSF	0.0582	43.1	0.03984	14.8	4.9	557	615
	FO-FS	0.1173	86.8	0.0368	13.6	0.328	37.3	139
	FO-FS / FO-FO-FS	0.1443	106.8	0.0365	13.5	0.257	29.2	150
PROPULSION	SINGLE	0	0	2.578	955.5	0.620	19.6	976
	NSF	0.0671	49.7	0.0401	14.9	8.17	258.2	323
	FO-FS	0.1382	102.3	0.03548	13.1	0.201	6.4	122
	FO-FS / FO-FO-FS	0.1787	132.2	0.03534	13.1	0.129	4.1	140
COMMAND AND CONTROL	SINGLE	0	0	18.33	6,794	—	—	6,794
	NSF	0.1901	140.7	0.4768	176.7	—	—	317
	FO-FS	0.3515	180.1	0.4509	187.1	—	—	427
	FO-FS / FO-FO-FS	0.4510	333.7	0.4498	166.8	—	—	500

The maintenance penalty was obtained by determining the cost of replacing a module in the RNS vehicle in earth orbit if a maintenance operation were required. The result showed that a Class 1 propellant module would cost $128.8 million to replace and a propulsion module would cost $31.6 million. These values were multiplied by the expected number of replacement modules for each criterion (Table 3) to arrive at the maintenance penalty.

The total cost for a criterion is then the total of the three costs.

The Class 3 weight, reliability, and maintenance penalties were calculated in a similar manner, but eight propellant modules are considered instead of the one module used on the Class 1 RNS. The Class 3 results are shown in Table 4.

Table 4
CLASS 3 RNS

MODULE	CRITERIA	WEIGHT EQUIVALENT VEHICLES	WEIGHT PENALTY ($10⁶)	MISSION FAILURE EQUIVALENT VEHICLES	MISSION FAILURE PENALTY ($10⁶)	MAINTENANCE EQUIVALENT MODULES	MAINTENANCE PENALTY ($10⁶)	OVERALL TOTAL COST ($10⁶)
PROPELLANT	SINGLE	0	0	6.84	2,558	0	0	2,558
	NSF	0.535	350	0.2911	77.3	26.3	265	692
	FO-FS	1.069	700	0.2741	72.7	1.69	17.1	790
	FO-FS / FO-FO-FS	1.322	865	0.2735	72.6	1.33	13.4	851
PROPULSION	SINGLE	0	0	2.5776	684.1	0.62	19.8	704
	NSF	0.0785	51.3	0.0401	10.7	8.17	258.2	320
	FO-FS	0.1617	105.8	0.0355	9.4	0.201	6.4	122
	FO-FS / FO-FO-FS	0.2091	136.8	0.0354	9.4	0.129	4.1	151
COMMAND AND CONTROL	SINGLE	0	0	18.33	4,864	..	..	4,864
	NSF	0.1901	145.6	0.4768	126.5	..	..	272
	FO-FS	0.3515	269.3	0.4509	119.7	..	..	389
	FO-FS / FO-FO-FS	0.4510	345.5	0.4498	119.4	..	..	465

Sensitivity Analysis

The effect of possibly inaccurate estimates of weight, reliability, maintenance requirements, and the dollar costs of these was investigated by calculating a new overall cost of a design criterion with these factors varied by a factor of four. The procedure used was to vary the total cost associated with the added weight (Tables 3 and 4), the mission failures, and the maintenance requirements by a factor of 2 and 0.5 individually. Altogether, 25 cases were run to represent the total number of combinations of varying the three parameters by the factor of four. The results should give the overall costs for each candidate design criterion for each module if the original estimates of the weight, reliability, maintenance requirements, or dollar costs were wrong by a factor of two, either high or low.

Results

The results of this specific trade study using current component failure rates are given in Tables 3 and 4. Tables 5 and 6 give the results for anticipated improvements in the state of the art. The single greatest contribution to the unreliability is caused by leakage. Tables 5 and 6 reflect specific design solutions (i.e., welded flanges) to achieve the higher reliability. The results of Tables 5 and 6 are considered the baseline. The optimum criterion is the same for both assumptions. Tables 3 and 5 show that the FO-FS criterion is optimum for the Class 1 propellant module, reflecting the high maintenance penalties (high cost of replacing a module using a Saturn V launch vehicle). The establishment of the FO-FS criterion will save about $560 million over the life of the program. Table 3 also shows that increasing the design complexity to FO-FS/FO-FO-FS will be detrimental to the amount of $12 million.

Table 5
CLASS 1 RNS

MODULE	CRITERIA	WEIGHT EQUIVALENT VEHICLES	WEIGHT PENALTY ($$10^6$)	MISSION FAILURE EQUIVALENT VEHICLES	MISSION FAILURE PENALTY ($$10^6$)	MAINTENANCE EQUIVALENT MODULES	MAINTENANCE PENALTY ($$10^6$)	OVERALL TOTAL COST ($$10^6$)
PROPELLANT	SINGLE	0	0	0.943	349.6	0	0	350
	NSF	0.0582	43.1	0.01992	7.4	2.5	279	330
	FO-FS	0.1173	86.8	0.0183	6.8	0.164	18.7	112
	FO-FS / FO-FO-FS	0.1443	106.8	0.0182	6.8	0.129	14.6	128
PROPULSION	SINGLE	0	0	1.289	477.8	0.310	9.8	488
	NSF	0.0671	49.7	0.02	7.5	4.09	128.1	186
	FO-FS	0.1382	102.3	0.01773	6.6	0.100	3.2	112
	FO-FS / FO-FO-FS	0.1787	132.2	0.01767	6.6	0.065	2.1	141
COMMAND AND CONTROL	SINGLE	0	0	9.16	3,397	--	--	3,397
	NSF	0.1901	140.7	0.2383	88.4	--	--	229
	FO-FS	0.3515	260.1	0.2254	83.6	--	--	344
	FO-FS / FO-FO-FS	0.4510	333.7	0.2249	83.4	--	--	417

Table 6
CLASS 3 RNS

MODULE	CRITERIA	WEIGHT EQUIVALENT VEHICLES	WEIGHT PENALTY ($$10^6$)	MISSION FAILURE EQUIVALENT VEHICLES	MISSION FAILURE PENALTY ($$10^6$)	MAINTENANCE EQUIVALENT MODULES	MAINTENANCE PENALTY ($$10^6$)	OVERALL TOTAL COST ($$10^6$)
PROPELLANT	SINGLE	0	0	4.82	1,270	0	0	1,279
	NSF	0.535	350	0.1456	38.7	13.2	133	522
	FO-FS	1.069	700	0.1371	36.4	0.85	8.6	745
	FO-FS / FO-FO-FS	1.322	865	0.1368	36.3	0.67	8.7	908
PROPULSION	SINGLE	0	0	1.2888	342.1	0.31	9.8	353
	NSF	0.0785	51.3	0.02	5.4	4.09	129.1	186
	FO-FS	0.1617	105.8	0.0178	4.7	0.1	3.2	114
	FO-FS / FO-FO-FS	0.2091	136.8	0.0177	4.7	0.065	2.1	144
COMMAND AND CONTROL	SINGLE	0	0	6.17	2,432	--	--	2,432
	NSF	0.1901	145.6	0.2383	63.3	--	--	209
	FO-FS	0.3515	269.3	0.2255	58.9	--	--	329
	FO-FS / FO-FO-FS	0.451	345.5	0.2249	58.7	--	--	405

The results for the Class 3 propellant module (Tables 4 and 6) are not the same due to the difference in cost of replacing a module (using the less expensive space shuttle) and the higher cost of adding components (each added component is added to eight modules). This latter effect can be seen by comparing the cost of increased weight for the Class 3 propellant module for the FO-FS criterion ($700 million) versus the cost for the Class 1 propellant module ($86.8 million). The optimum criterion, therefore, for the Class 3 propellant module is NSF.

The sensitivity analysis was designed to demonstrate the strength of these results. In the case of the Class 1 propellant module, the FO-FS criterion remained optimum in 22 of the 25 cases when the several factors were varied by a factor of 4. In three cases the optimum criterion changed to FO-FS/FO-FS-FS when the maintenance costs were increased by a factor of 2, reflecting even a greater need for redundancy. For the case of the Class 3 propellant module, the sensitivity analysis results show that the optimum criterion changed nine times out of the 25 cases, reflecting a less firm conclusion. In these nine cases the criterion changed from NSF to FO-FS when the cost of increasing the redundancy (weight penalty) was reduced by a factor of 2.

The FO-FS criterion is optimum (Tables 5 and 6) for the propulsion module for both RNS concepts. This selection did not change in the 25 cases run in the sensitivity analysis and thus indicates a firm selection. The FO-FS criterion compared with a single thread design will save several hundred million dollars over the life of the program and reflects the high maintenance and mission failure penalties. The high mission failure penalty for the Class 3 propulsion module ($64.8 million) acts against the single failure criterion, and the large maintenance penalty ($258 million) acts against the NSF criterion.

The lack of a maintenance penalty for the command and control module, due to the ground rule that the module is returned to the surface of the earth for refueling after every mission, dictates that the NSF criterion is optimum. This is true for both concepts and is a strong selection as reflected by the sensitivity analysis where all 25 cases selected this criterion.

The results discussed above were for a specific project with specific ground rules. It is apparent that the selection of the ground rules dramatically affects the results, as demonstrated in the case of the command and control module. The specific results of this trade study are therefore not applicable to any other project, but the insight gained and some general conclusions can be.

Figure 5 is a plot of the mission loss probability (1 minus Reliability) and the maintenance requirements versus system weight as components were added. This figure represents the entire Class 1 RNS. Figure 5 vividly illustrates the phenomenon that as components are added starting with the single thread design the mission loss probability decreases sharply, but the maintenance requirements increase due to the larger number of components that can fail. After the NSF design point, adding components causes the maintenance requirements to drop dramatically, but the mission loss probability decreases only slightly. The drop in maintenance requirements is due to the "dispatch with inoperative equipment" rule (if only one out of three components fail, the RNS can leave earth orbit without maintenance). The reliability does not increase significantly because the large reliability benefits are derived in going from one component (single thread criterion) to two redundant components (NSF). The addition of a third component (FO-FS) does not significantly increase the reliability and can reduce it due to failure modes which are not reduced by redundancy (external leakage of a valve).

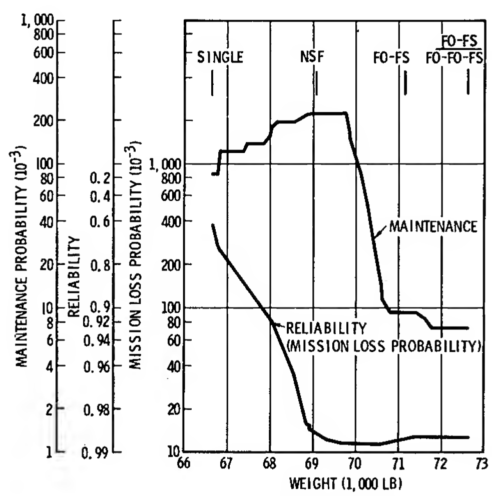


Figure 5. Reliability and Maintenance Requirements versus Weight, Class 1-H RNS

The following conclusions can be drawn: the single thread criterion is only applicable where weight or initial cost is paramount, and reliability and maintenance are not important problems; the NSF criterion is very efficient from a mission success (reliability) standpoint but produces high maintenance requirements; the FO-FS criterion is very efficient from a maintenance standpoint but does not provide any appreciable increase in mission success and in fact can reduce the mission success probability; and the FO-FO-FS criterion does not produce significant improvements in reliability or maintenance. The

optimum criterion therefore depends on the relative importance of weight, reliability, and maintenance.

The study shows that for projects such as the RNS, space shuttle, commercial aircraft, etc. where the requirement for maintenance could seriously affect the operation schedule, the FO-FS criterion is optimum. Projects such as an unmanned probe to Mars, a lunar landing mission, or any project where maintenance is not performed should use the NSF criterion. The single thread design should only be be used where mission failure does not introduce a safety problem (to a crew or the general public), no maintenance is required (or maintenance can be performed inexpensively, easily, and does not affect operations), and weight is the major factor.

Summary

The subject of design criteria has been raised on several recent space programs and specific criteria imposed on some of them. An interesting and vital question is usually asked when this happens, i.e., which criteria should be applied and why? On one project, the Reusable Nuclear Shuttle, a study was conducted to answer this question. The study investigated the overall program cost that was incurred when each of four design criteria was applied. This included the cost of the additional components to achieve the several criteria, the cost of mission failures, and the cost of maintenance. The lowest overall program cost was selected as optimum. The insight gained and the general conclusion developed in this specific trade study are applicable to other aerospace projects.

References

[1] Boehmer, C. B., "The Reusable Nuclear Shuttle—Its Operation, Test Program, and Problem Areas", Institute of Environmental Sciences, April 1971.

[2] Boehmer, C. B. and Cowls, R. S., "System Aspects of a Reusable Interorbital Nuclear Shuttle", American Nuclear Society 16th Annual Meeting, June 1970.

[3] Riedesel, R. G. and Cowls, R. S., "Manned Planetary Missions with Reusable Nuclear Shuttle", American Astronautical Society 16th Annual Meeting, June 1970.

[4] Nuclear Shuttle System Definition Study—Phase III, Final Report, MDC G2134, May 1971.

AN INTEGRATED APPROACH TO HARDNESS ASSURANCE*

L. J. Palkuti
Naval Research Laboratory
Washington, D. C.

R. K. Thatcher
Battelle, Columbus Laboratories
Columbus, Ohio

Introduction

Just as the other assurance sciences of reliability, quality, maintenance, and safety independently grew up with their own technology, so has radiation hardness. It is now time, however, for hardness assurance to be integrated with the other assurance sciences, to compete equally for attention and action, and to interact with the other assurance sciences to achieve an optimum blend of system characteristics which will satisfy the total system specifications. As nuclear hardness specifications become more prevalent, it becomes necessary to address more openly the problems associated with assuring that each item produced will meet its radiation specifications.

It is the purpose of this paper to introduce hardness assurance to you in such a way as to make its operational procedures familiar and easily integrable with your present system assurance programs. Because of the constraints on the extent of this paper, no attempt was made at discussing "how-to-do" specific hardness assurance (H/A) efforts.

Discussion

Background

To put H/A in proper perspective, it should be recognized that H/A is but one facet of a total system assurance program--system assurance being defined to include all planning, control, monitoring, and evaluation efforts related to system reliability, quality, maintenance, safety, and nuclear hardness. As shown in Figure 1, we are considering there to be five major functions associated with achieving the successful establishment of a system. They are concepts and specifications, engineering, production, system assurance, and deployed systems under maintenance. Specifically in this paper only, the H/A effort, which is under the control of the manufacturer, will be discussed. However, it should be noticed that this H/A effort as defined still can and will have to interact outside the manufacturing environment with those persons involved in definition of concepts and specifications and those performing system maintenance. In addition, the organization as outlined in Figure 1 also suggests that the manager of the system assurance group be responsible only to the program manager who can arbitrate the differences between the system assurance, engineering, and production groups. The influence of system assurance should be equal to that of engineering and production. In this way, H/A like all other assurance sciences will receive appropriate attention. In addition, the integration of H/A with the other assurance sciences results in the application of the concepts, principles, and tools already developed for reliability, quality control, maintenance, etc., for effecting a cost-effective overall system assurance program.

The primary emphasis of H/A on a particular system is placed on electronics, radiation shields, and electromagnetic shields. The efforts associated with electronics H/A are keyed to assuring that only those parts are incorporated into the system that meet the required radiation resistance measure and that the radiation resistance is maintained through all higher levels of assembly. The work associated with shielding H/A consists of maintaining the proper protection for the electronic equipment. This includes efforts to assure that each portion of the overall shield will maintain the appropriate amount of attenuation and that proper installation of shield components results in no gaps or discontinuities.

A group of H/A tasks are defined to effectively carry out a H/A effort. They are policy formulation, product analysis, process analysis, product and process control, testing, statistical methods, failure analysis/product modification, and system analysis. Selected groups of these tasks are utilized at 5 discrete phases during the manufacture of the system. These 5 discrete phases are

- Engineering
  - DESIGN
- Production
  - INCOMING PARTS
  - ASSEMBLY PROCESSES
- System Assurance
  - SPECIAL PROBLEMS
  - EVALUATIONS.

Task assignment at these five phases during manufacturing are considered important for two reasons:

(1) To some extent in all H/A efforts and in particular, on large, state-of-the-art systems, slightly different backgrounds and viewpoints are needed at each phase. For example, in the area of incoming parts, a system pushing the state of the art in the radiation resistance of semiconductor parts would require a person with <u>detailed</u> background in how semiconductor parts are presently being made. On the other hand, the assembly processes for the same system would require a person who understood the influence of soldering, human handling, circuit board cleaning, etc., on reliability, hardness, etc. Most likely this would be two different people.

(2) These phases of manufacturing also happen to have associated with them discrete (but not independent) H/A documentation.

It is important to note that these phases are easily identified in all systems and act as a focal point

*Original work performed under Contract DASA-01-69-C-0116.

for the tasks (the "what-to-do") of H/A. Because they do act as focal points, the H/A effort, regardless of system or level of radiation specified, has to address to some extent the problems associated with each of these phases of manufacturing. The results of addressing these problems needs to be documented clearly and concisely. More discussion of these phases are presented in synopsis form below.

Design

Since the total H/A program is oriented toward a specific system, each task within the total H/A program must be tailored to the specific system requirements. As a result, the design assurance activities--interacting directly with the design process--establish a specific H/A plan and develop specific H/A requirements. In addition, this phase includes effort to eliminate or locate possible H/A-related problems before the start of formal production. To achieve this, the performance requirements and radiation specifications of the system must be translated to be the hardness requirements of specific parts and components.

The H/A tasks applicable to the design phase include formulation of H/A policy, product analysis, and process analysis. Some specific applications of these tasks involve charting the production process; the analysis of function and circuit requirements; definition of part requirements and tolerances; establishment of experimental methods and correlation with the radiation specification; classification of the variability of the radiation response of components from development tests; establishment of the variability of specific processing methods; and component evaluation.

The design assurance tasks can be carried out by any group with sufficient background in all areas to be covered, the important emphasis being that the group has broad insight into the radiation response capabilities of components and assembled systems. They certainly should consult with other persons who may later become involved in the H/A effort. In any situation, these design assurance tasks should be performed to the fullest extent on all equipment with a radiation specification.

The end product of the design phase activity should be a clear definition of system requirements, including the required hardness of critical parts and assemblies with respect to the various specified nuclear radiations, measured in terms of product and process parameters. In addition, documentation should be prepared which provides a clear definition of the various constraints and assumptions (design trade-offs) that lead to the specification of the methods and procedures needed to control, monitor, and evaluate the various components and assemblies of the system. Also included in the documentation would be methods for initiation of vulnerability information feedback and corrective action.

Incoming Parts

The scope of incoming parts assurance covers all H/A activities associated with the qualification of part suppliers, the evaluation of material and processing procedures used in the manufacture of the parts, and the actual acceptance and evaluation of incoming purchased material and parts. Included are those parts and materials received from outside sources and those produced by other plants of the same company or other divisions of the same plant. Since the development of radiation hard parts generally precedes the production contract, incoming parts assurance activities are performed during both the development and production of the system.

The H/A tasks applicable to incoming parts assurance include process analysis, product and process control, radiation testing (sampling), and failure analysis. Some specific applications of these tasks involve vendor capability evaluations and survey of vendor facilities; clear delineation of part requirements and specification; vendor qualification and certification of material and parts; vendor process control and inspection procedures; vendor testing including preconditioning and screening of parts; selection of proper feedback and corrective action procedures; implementation of buyer screens and lot-sampling plans; and development of parts radiation tests including data analysis.

The tasks associated with this effort should be done primarily by the manufacturer's own assurance technology personnel. Some assistance could be solicited elsewhere to provide insight to specific part and material manufacturing processes, process controls, and screens used by the supplier.

These tasks can be carried out to varying degrees depending on the radiation levels specified and the ability of the parts/materials to perform their specified function in the system. It is possible that no assurance work beyond what is to be applied to incoming parts/materials is necessary.

The documentation associated with this effort should be

(1) Those portions of the material and part procurement specifications which relate directly to hardness assurance

(2) Guidelines to the procurement office as to "one-time allowances" or, at least, a procedure to follow if the manufacturer desires to ship parts/materials which do not meet in all aspects the procurement specification

(3) Detailed procedures to be utilized in testing and screening incoming parts including acceptance and rejection criteria

(4) The baseline processing document prepared by the part manufacturer and which as been evaluated and corrected.

Assembly Processes

As design assurance is oriented toward specific design features of a system, assembly processes are oriented toward specific processing, assembling, and packaging features of the system. Of particular concern are the ways in which these production features impact the hardness of modules, circuits, subsystems, and the complete system. Specifically, assembly assurance involves the control of the radiation hardness at the source of production and throughout deployment so that departures from hardness specifications can be corrected before too soft systems are produced and so that the hardness of the system is maintained after deployment.

H/A tasks applicable to assembly assurance include process analysis, product and process control, radiation (sampling) tests, and failure analysis. Some specific applications involve analysis of specific process

capabilities and requirements which impact radiation hardness; institution of process controls on assembly and packaging techniques; screening procedures and lot-sampling radiation tests; maintenance control on deployed systems; and tests of parts, modules, and systems returned from the field including analysis of operational experience and maintainability experience.

The tasks associated with assembly assurance should be done primarily by the manufacturer's own assurance technology personnel. Some assistance could be solicited from the people involved in maintenance of the equipment.

The appropriate tasks will be carried out to varying degrees depending on the radiation levels specified and on the ability of the parts/materials to perform their specified function in the system. It is very likely that no special activity will be performed. Usually, EMP shields will have to have some controls placed on them, and some guidelines for maintenance of shields will have to be prepared.

Documentation related to this effort will include a complete description of the manufacturing processes, testing, screening and rework cycles necessary to assure the hardness of the product. In addition, maintenance practices which may have an impact on the hardness of the system should be defined and the appropriate procedures clearly specified.

Special Studies

In the course of system production, problems reflecting on the hardness of the system can occur that require a concentrated and rapid solution. Typical of such problems are a sudden increase in incoming part failures, a sudden increase in the failures observed in a module screening procedures, or an observed anomaly in the radiation test results of a module. Special studies then need to be implemented to locate the causes of deterioration of the hardness of either parts or assemblies and to determine the possibilities for improvements in the general H/A program. These studies are directed toward major, usually nonrepetitive, problems requiring activity from either of several groups in the contractor organization or a cooperative effort between the contractor and a specific vendor or other organizations.[1] Special-study activities are not only initiated to find a solution for hardness-related troubles experienced during production but also to conduct major investigations for either developing new or improving present techniques of attaining hardness standards of a particular product or process.

The techniques used in special studies consist largely of special applications of the standard methods used in the other tasks of hardness assurance. Some of these techniques--which are common to all assurance tasks--are special chemical and physical analysis methods; detailed and specialized failure mode determination; and specialized statistical methods. The fundamental features of these studies are (1) the coordination of effort so as to utilize all available resources in an integrated approach to the specific problem and (2) the use of the best technical methods in conjunction with a technologically sound approach to achieve a solution whose presence or lack of impact on hardness is clearly understood.[2]

This activity can be performed by any group with the specialized capabilities required to address the problem. The important aspect of the performance is to get the correct diagnosis and most economical solution, even if the work has to be done outside the plant. The manager should be aware of both the capabilities of in-house and outside contractors who can do this type of work. A relationship should be established with both in order to facilitate a fast response to the problem.

The extent of this type of activity may turn out to be minimal, but in planning, several such studies should be budgeted even for the low-level specification.

A report containing as a minimum: statement of the problem, approach to solution, statement of causes for the problem, and solution(s) which will alleviate the problem. Typically, the problem solution would be in terms such that it can be translated easily to implementation procedures by appropriate H/A task personnel.

Evaluation

Assessment of the effectiveness of the control phases of the H/A program through appraisal of the achieved hardness of the product comprises hardness evaluation. By testing and analysis, the probability of system failure to withstand required stresses (at least, up to credible stresses of interest) is generated. Since it is generally impractical to test enough systems to generate the data base that would provide mathematical confidence in system hardness, evaluation must utilize a mixture of statistical data (usually available from parts, module, subsystem, and occasional system tests), analysis, and engineering judgment based on understanding of the important system failure modes.[2]

The H/A tasks applicable to hardness evaluation include the analyses of numerous types of test data, failure analysis, and system hardness analysis. Some specific applications of these tasks involve selection, storage, and retrieval of data generated by the screening and catastrophic and degradation failure modes; analysis of specific component failures and appropriate evaluation of anomalous test results; incorporation of process and test modifications including planned experiments to evaluate the effects of process or test changes on system hardness; use of mathematical models to establish the relation between part and module response and the system response; and mathematical simulation of the inspection system to determine the effectiveness of controls, screens, and lot-sampling procedures.

The tasks related to evaluation can be carried out by any group or groups with sufficient background in testing and analysis. The important aspects of these efforts are that the work be performed consistently, that the tests be carried out as close to actual operating conditions as possible, and that rapid and complete feedback to the other H/A tasks be maintained.

The extent of performance of these tasks depends on the radiation specifications and the engineering judgment applied to the problem. In some cases, it could be minimal.

Since the primary goal of these tasks is to obtain the probability of product survival and the associated confidence level, the documentation should describe procedures for obtaining them from the data generated. The documentation should also contain the data generated and the complete details for generating the necessary data.

Hardness Assurance Tasks

In order to better appreciate the scope of the assurance tasks, a brief discussion of the more important tasks is presented.

Basic to the implementation of each of these tasks is the following set of steps:

- Determination of the need and scope of the task activities
- Development of procedures for and the framework within which the task is conducted.
- Establishment of methods for systematically evaluating the effectiveness of the task procedures and for properly instituting corrective measures
- Documentation of the rationale, procedures, methods, and results as the task is performed.

The tasks to be discussed are Formulation of Policy, Product Analysis, Process Analysis, Product and Process Control, Testing, Statistical Methods, Failure Analysis/Product Modification, and System (Hardness) Analysis.

Formulation of Policy

Policy formulation is reflected throughout all of the manufacturing phases even though it is primarily applied during design. This is clearly one task which benefits from previous experience and extensive insight into the state of the art of radiation effects. The requirements for any H/A analysis and implementation is a clear delineation of the objectives of the H/A effort. To formulate policy, the assumptions and constraints within which the H/A effort is conducted must be clearly understood and stated. Past experience and the baseline study in a system development program are used as a guide to provide the basic understanding of the relevant vulnerability modes. This understanding will lead to valid analytical and experimental methods to establish confidence in hardness by providing answers to four general questions:[3]

(1) Has the design adequately corrected for the identified failure modes?

(2) What are the implications of the allowed variations in characteristics of the elements of which the system is composed?

(3) Are there any significant unidentified failure modes?

(4) What is the effect of aging and handling on system hardness?

The planning required to be sure that these questions are properly addressed requires (1) identification of the H/A decisions that must be made, (2) identification of the H/A problems that must be solved, and (3) specific documentation of the H/A policy.[2]

The H/A policy statements should exist at all times. As new information is obtained they can be extended and updated to reflect the latest understanding and methodology to implement a particular H/A policy. In addition, the documentation of the particular H/A policy should be continuously reviewed by experts who understand thoroughly how each H/A policy impacts system hardness and performance.

Product Analysis

This task also is a primarily utilized during the design phase. Determination of the need and scope of control and monitoring activities requires analysis of the factors that bear on the hardness of the production hardware. The act of analyzing involves breaking down the system into its elements and then synthesizing these elements back to the whole. The basis for the analysis is the set of radiation levels that constitute the radiation specification. This specification--along with the system design--and the addition of minimum system failure probability and associated confidence level quantify the required degree of product and process control. Normally, the failure probability is partitioned between the major failure modes associated with the product. Table 1 lists the various components of the radiation environment, the important failure modes, and their effect on electronic systems. The possible existence of secondary failure modes (i.e., failure modes occurring at radiation levels above that associated with the primary mode) must also be investigated. The partitioning of failure modes allows the delineation of the hardness requirements of the subsystems and modules, continuing logically to the requirements on circuits and components parts. For this to be done intelligently requires a thorough knowledge of the system and how each failure mode affects its function. Setting of module and device requirements is thus best accomplished by the product-design engineer (who should be most knowledgeable in the effect of radiation on the components, whose past experience with similar products can point out areas in design that can lead to hardness-related problems.

The previous considerations will result in a set of definitive criteria for radiation-critical part tolerances. The following examples illustrate different ways in which the tolerance might be specified. A simple logic gate will fail when the neutron fluence is sufficient to degrade the output transistor gain until it can no longer accept the "sink" current required at the specified fan-out. In this case, the minimum gain allowable at the criteria level is the critical factor. The neutron associated tolerance of a differential amplifier can be used as a different example. In a neutron environment perhaps the production of an excessive mismatch in the gain of a pair of transistors would be a more important factor in circuit failure than would be the absolute change in gain. Relative changes in gain could introduce (by modifying emitter crowding) changes in the input characteristics of the two transistors and result in an unacceptable offset.[4] For individual transistors the maximum allowable gain or the maximum saturation voltage may be the critical factors. Table 2 illustrates typical transistor requirements as determined from an analysis of a particular circuit.

Process Analysis

Process analysis is utilized in the design incoming part and assembly phases. This task complements product analysis. The features of the manufacturing processes used to fabricate the components, circuits, and the system are studied to determine the relationships between process parameters and the radiation response as well as indicate the process capability and process stability. The general approach to process analysis is to relate electrical and mechanical characteristics and the radiation sensitivity to process and material parameters. The physical parameters associated with the fabrication of the product can then be adjusted to satisfy electrical, mechanical, and radiation performance with maximum cost effectiveness. If the physical parameters can be related to

terminal measurements, then screening procedures can be incorporated directly in the electrical test plan such that the process parameters are maintained at predetermined tolerances. In case the radiation response is not directly related to electrical parameters, then additional physical measurements are required to maintain the process within the required tolerances. Typical of some of the material which might be generated during this phase of the analysis are the examples shown in Tables 3a and 3b.

Two useful techniques used in process analysis are pilot productions and specifically fabricated test vehicles. Test vehicles are especially useful in the process analysis of complex integrated circuits where requirements include new or special processing. The use of test vehicles involves the fabrication of test patterns or test elements that are specifically designed to permit direct measurement of all critical process parameters at different stages of fabrication. In addition, these devices also furnish an indication of device radiation sensitivity and quality. Test vehicles can either be utilized before die scribing (indicating base and collector resistivity, base profile, thin-film resistivity, contact resistance, etc.) or they can be assembled following wafer probe and chip scribing and subjected to electrical characterization and environmental testing. Some examples of certain test devices used to evaluate processing include transistor elements, capacitors, and metallization and resistor elements.

Product and Process Control

Product and process control is primarily applied during the incoming parts and assembly phases of manufacturing. After the requirements of the product and processes have been established, implementation requires a program of product and process control applied throughout all critical production steps. The scope of control activities (related to a H/A program) for the typical hardened electronic system involves direct control of materials, parts, components, and assemblies throughout the production cycle. In particular, to ensure that the system hardness objectives are met, control is applied in three major areas of system production including semiconductor parts, module and circuit assembly, and assembly of shielding enclosures. Since semiconductor devices provide the foundation of the radiation hardness of electronic systems much effort is expended in the control and screening of these devices. Some specific techniques which are used in product and process control are

(1) Use of test vehicles

(2) Baseline process control

(3) Variables data monitoring

(4) Screening

(5) Lot sampling.

Specifically fabricated test vehicles can be useful in process analysis to evaluate radiation sensitivity as well as reliability. Documenting the process using baseline specifications is important for maintaining consistent hardness. An example of the process details typically covered by the "baseline" are listed in Figure 2.

The usefulness of continued monitoring of key variables data is illustrated in Figure 3. Process variations can be quickly observed and appropriate corrective action taken.

The need for device screening is illustrated in Figure 4. The variation of radiation hardness shown necessitates continued device screening. The technique of irradiate-anneal screening is useful when the possibility of maverick devices exists. Such a technique with the associated reliability data is shown in Figure 5.

The more important control and screening techniques for neutrons and ionization effects are presented in Tables 4 and 5, respectively.

Testing

The technique of radiation testing (sampling) provides direct experimental verification that (1) the control procedures utilized during production processes are adequate to ensure the previously defined standards and (2) the appropriate level of radiation hardness has been achieved throughout the useful life of the equipment. These radiation tests are generally conducted on a sampling basis, i.e., a portion of production output including potential parts, circuits, modules, subsystems, and systems is tested and the remainder is accepted, modified, or rejected according to the results of these sampling tests. Sampling radiation tests represent a supplement to routine inspection, electrical testing, in-process controls, and screens. They provide a direct measure of the test-item response made during or after exposure to simulated radiation environments that can be related to system hardness capability.

Test Rationale

A well structured program will contain testing for each effect of radiation that has an impact on the system performance. Some of the available simulation techniques are[2]

(1) For prompt pulse ionization effects; linear accelerators, flash X-ray machines, and underground nuclear tests.

(2) For delayed ionization effects: pulsed nuclear reactors (TRICA, SPR), linear accelerators, flash X-ray machines, and Cobalt-60 sources.

(3) For neutron-produced displacement effects: steady-state reactors and pulsed reactors.

(4) For short-term annealing studies of displacement effects: pulsed reactors.

(5) For EMP: large, parallel-plate high-voltage facilities such as the ALECS facility at AFSWC. Since the EMP tests must be performed at the system level such facilities must be large. These tests can be supplemented by current-injection experiments and scaled-intensity, full-scale-dimensions tests such as on long-wire EMP simulators.

The nuclear specification never exactly represents a nuclear threat because of the variety of relevant environments within the specification envelope and because all simulation facilities fail to provide an exact simulation of the threat environment. This fact makes realistic simulation of the operational environment impossible.[2]

It is particularly fallacious to assume that the specification envelope represents the worst case. There are numerous examples in which a system will fail at an exposure level inside the "survivability envelope" but which will survive at the envelope exposure level because of the a compensating correction mode.

System (Hardness) Analysis

Each test conducted under the H/A program provides additional data on the hardness capability or survivability of the system. The standard method of design to meet a requirement followed by proof test does not, in general, provide sufficient statistics to determine analytically the capability of the hardware poplulation to meet the requirement. But the continued acquisition of test data by hardness assurance testing does provide an ever expanding data base which yields an ever improving estimate of the system survivability--derived with statistical confidence.[14] The systematic translation of this data into estimates of system hardness requires specialized techniques of data storage and statistical analysis. The relationship between production flow, H/A-related testing, and statistical system analysis is illustrated in Figure 6.

The hardness of a population of production systems differs due to variability of part radiation response and variation of manufacturing processes. From an understanding of the structure of the system and the radiation response of components from all levels of assembly, an estimate of the system hardness is made. The flow of information from these test results through system analysis is illustrated in Figure 7. Utilizing known failure modes, a preselected failure budget, statistical tests, and engineering judgment, a system-hardness model is chosen. Statistical methods may be used to compute the parameters of the model. The results may be presented as analytical expressions, such as failure density functions or hazard rates and their associated confidence and tolerance

The system hardness from catastrophic failure modes can be estimated chiefly from failure distribution of parts data. Specific methods from system reliability analysis can be studied to determine suitable procedures. The system hardness from degradation failure modes is estimated in a similar manner but with emphasis on assembly level rather than parts data.

Conclusions

The manner and the extent to which the H/A tasks are performed, i.e., "the how-to-do-it", is of course, a function of the system being produced. The specific implementation of the tasks must be keyed to the constraints placed on the total system. Some factors that influence H/A decisions include the radiation environmental criteria, complexity and difficulty of the mission, the flexibility of the design, complexity of the system, reliability expected of the system, priority placed on support for reliability and for hardness, production schedule, access of the system for design changes and maintenance, funds available, the extent and length of the study and preproduction phases of the system program, and the state of the art at the time H/A decisions are made. These factors have a dramatic effect on the methodology of the specific task requirements and thus on the overall structure of the H/A program. These factors, however, should not influence the quality of the program developed, i.e., the appreciation for performing tasks correctly and the understanding for the effective implementation of these tasks. Care must be exercised to avoid eliminating or weakening requirements to fit contractor capabilities or operating procedures.

In summary, the following points are important:

- Hardness assurance should be incorporated with the other assurance sciences; i.e., incorporated with reliability, quality control, maintenance, and safety.
- Hardness assurance as part of the other assurance sciences should be managed independently of engineering and production and should only be responsible, like engineering and production to the person responsible for the total program.
- The hardness assurance tasks should be addressed to some extent on all systems. They are applied on five points of interaction with manufacturing: (1) design, (2) incoming parts, (3) assembly processes, (4) special studies, and (5) hardness evaluation.
- The hardness assurance tasks must be tailored to fit uniquely the specific program being addressed. This means that the appropriate combination of assurance technologies is to be used only to the degree necessary.
- It is imperative that one man be given the responsibility to see that the work is carried out completely, because visibility of the hardness assurance efforts can become somewhat clouded when incorporated into a company's operating procedures.
- All documentation should be carefully reviewed to insure that it is complete, independent, and unambiguous. This is particularly important if subcontractors are involved.
- Keep in mind not only one's own requirements for H/A but also those under which the other person has performed his work, in reviewing H/A literature or selecting technologies developed by other groups.
- Care must be exercised to avoid eliminating or weakening requirements to fit capabilities or operating procedures.

References

(1) Feigenbaum, A. V., Total Quality Control, McGraw-Hill Co., Inc. 1961.

(2) van Lint, V.A.J., Gulf Radiation Technology, private communication.

(3) van Lint, V.A.J., "Confidence in Survivability", Gulf Radiation Technology Report, Gulf-RT-A10799 (August, 1971).

(4) Caldwell, R., The Boeing Co., private communication.

(5) "Lockheed Missiles and Space Company Semiconductor Controlled Line", unpublished.

(6) Sullivan, W. H., and Weinbin, J. H., "A New Propagation Delay Time Bridge for the Study of Displacement Damage Effects in Bipolar Transistors", IEEE Transactions on Nuclear Science, NS-18, 420-428 (December, 1971).

(7) Unwin, A. M., Arimura, I., and Rosenberg, C., "The Adequacy of Neutron Damage Normalization Techniques", IEEE Annual Conference on Nuclear and Space Radiation Effects (July 21-23, 1970).

(8) Bell, D. R., and Lueder, J. R., "Radiation Screening Techniques for Integrated Circuits", Lockheed Missiles and Space Co., AFWL-TR-70-12 for the Air Force Weapons Laboratory (August, 1970).

(9) Niehaus, D., Frank, M., Stoll, E., and Swick, E., "Neutron Hardness Assurance for Power Transistors", IEEE Transactions on Nuclear Scinece, NS-17 (6), 154-159 (December, 1970).

(10) "Nuclear Radiation Hardness Assurance for Electronics", Bendix Navigation and Control Division, Final Report to the Space and Missile Systems Organization (May 30, 1970).

(11) Unwin, A. M., and Rosenberg, C., "Improved Radiation Storage Time Predictions by a New Approach to Inverted Transistor Characterization", IEEE Transactions on Nuclear Science, NS-15 (6), 260-266 (December, 1968).

(12) Beadles, R. L., Simons, M., and Breen, W. M., "Integrated Circuit Visual Inspection Using Spatial Filtering", IEEE Transactions on Nuclear Science, NS-17 (6), 178-182 (December, 1970).

(13) "Hardness Assurance", Autonetics, Division of North American Rockwell Corp., BR 71-899/201 (October, 1971).

(14) Heaton, E. C., Rotolante, R. A., and Fitch, S. H., "Radiation Hardness Assurance for Electronic Systems", IEEE Annual Conference on Nuclear and Space Radiation Effects (July 21-23, 1970).

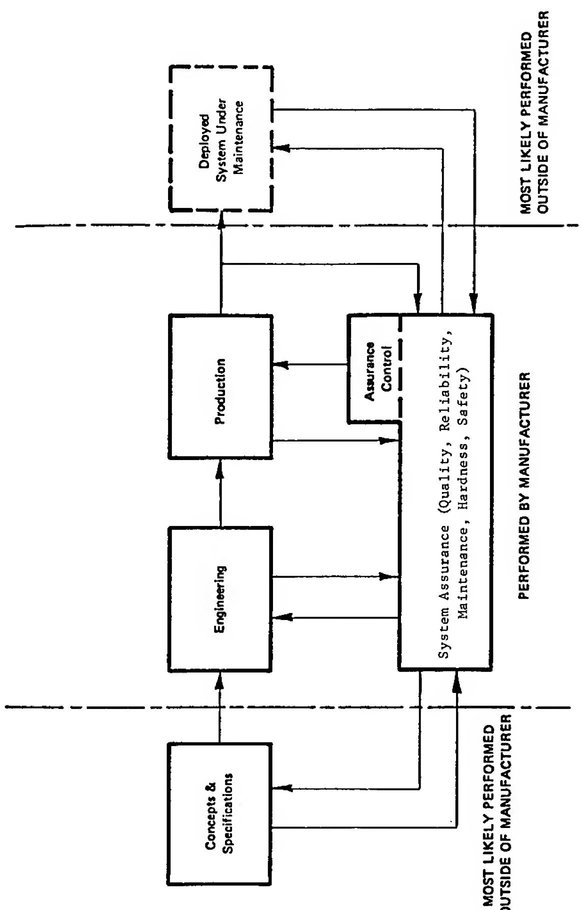


FIGURE 1. DIAGRAM OF THE PRIMARY DISCRETE FUNCTIONS AND THEIR INTERACTIONS ASSOCIATED WITH THE PRODUCTION OF A SYSTEM

Resistivity of Raw Si Wafer

- Measure

Initial Surface Preparation

- Prediffusion clean
- Oxidation

Diffusion Process

- Cleanliness controls
- Substrate temperature (furnace temperature)
- Impurity source density
- Diffusion times
- Sheet resistance (base and emitter)
- Bevel and staining (base diffusion depth)
- Mask inspection

Metalization

- Mask inspection
- Metalization deposition times
- Multiple metalization sources
- Evaporator control

- Silicon wafers are prepared by the fabricator using his normal processes and equipment working to specifications prepared by the fabricator.

FIGURE 2. EXAMPLE OF PROCESS DETAILS COVERED BY "BASELINE SPECIFICATION"

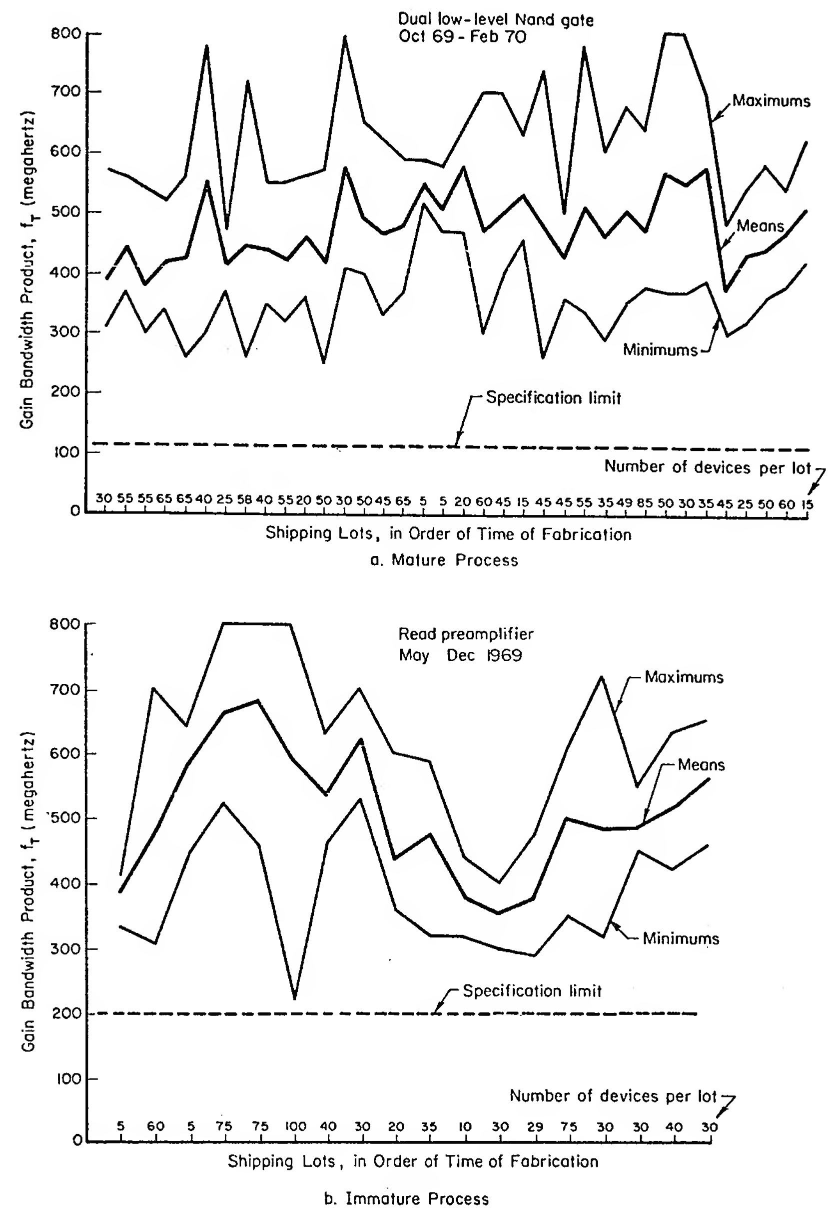


FIGURE 3. PROCESS VARIATIONS OVER TIME

RELIABILITY FAILURE RATES (200,000 DEVICE-HOUR LIFE TESTS)

Device Type	Test Condition	Percent Failure/1000 Hours[a], 90% confidence level: Recovered[b] From 5×10^5 rad	Control Group	JAN Mil Spec TX
2N1613		1.2	1.2	5.0
2N2411		1.2	1.2	5.0
2N1132		1.9[c] 1.2[d]	1.2	5.0
2N2219		1.9	1.9	5.0

(a) Under maximum rated power.
(b) Irradiated groups were "stress recovered" prior to life testing.
(c) For total group.
(d) For total group with nonradiation associated failure removed.

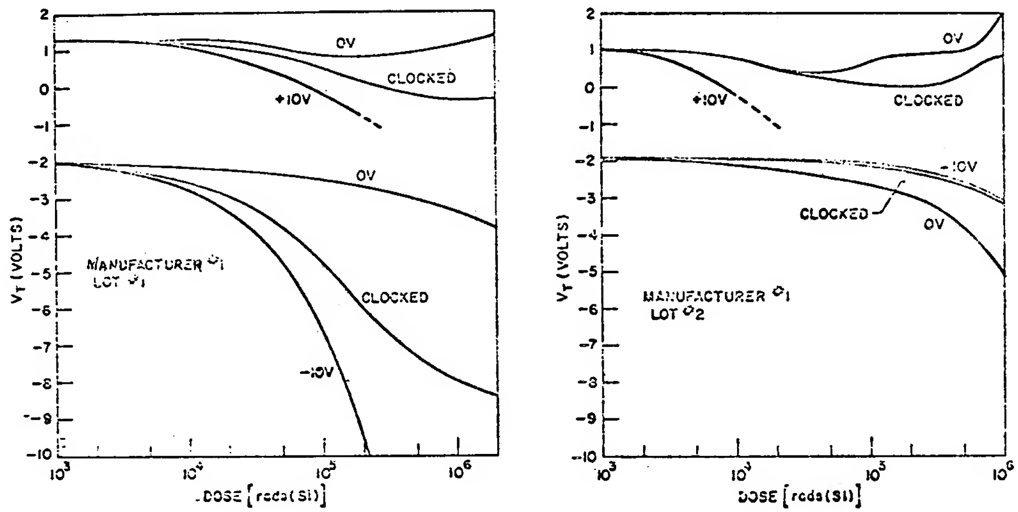


FIGURE 4. THRESHOLD VOLTAGE VERSUS TOTAL IONIZING DOSE FOR THE DC 4007A

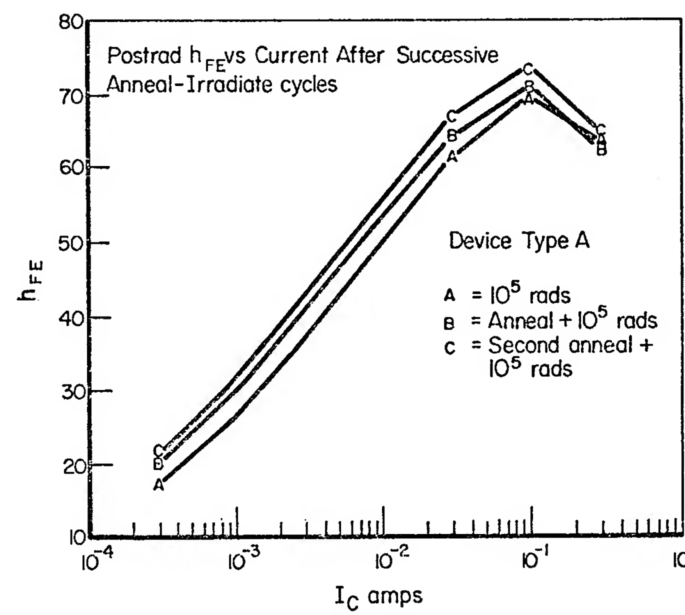


FIGURE 5. IRRADIATION-ANNEAL SCREENING

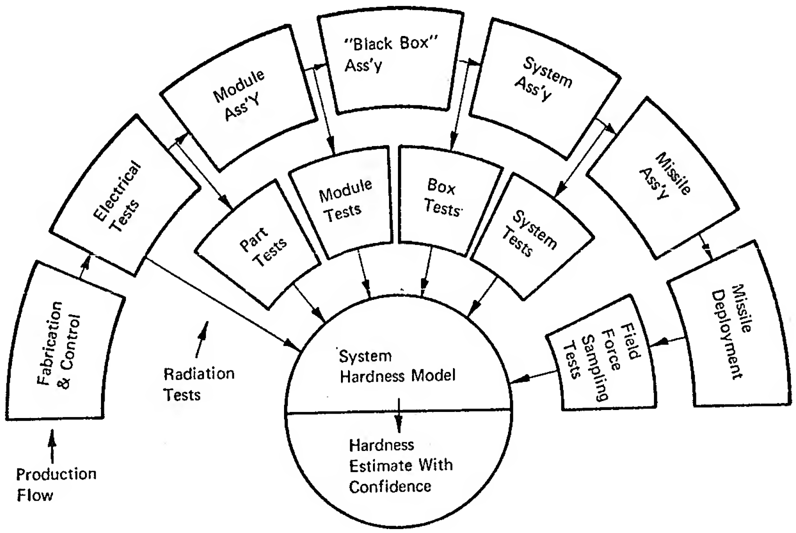


FIGURE 6. DATA INPUT FOR SYSTEM ANALYSIS (13)

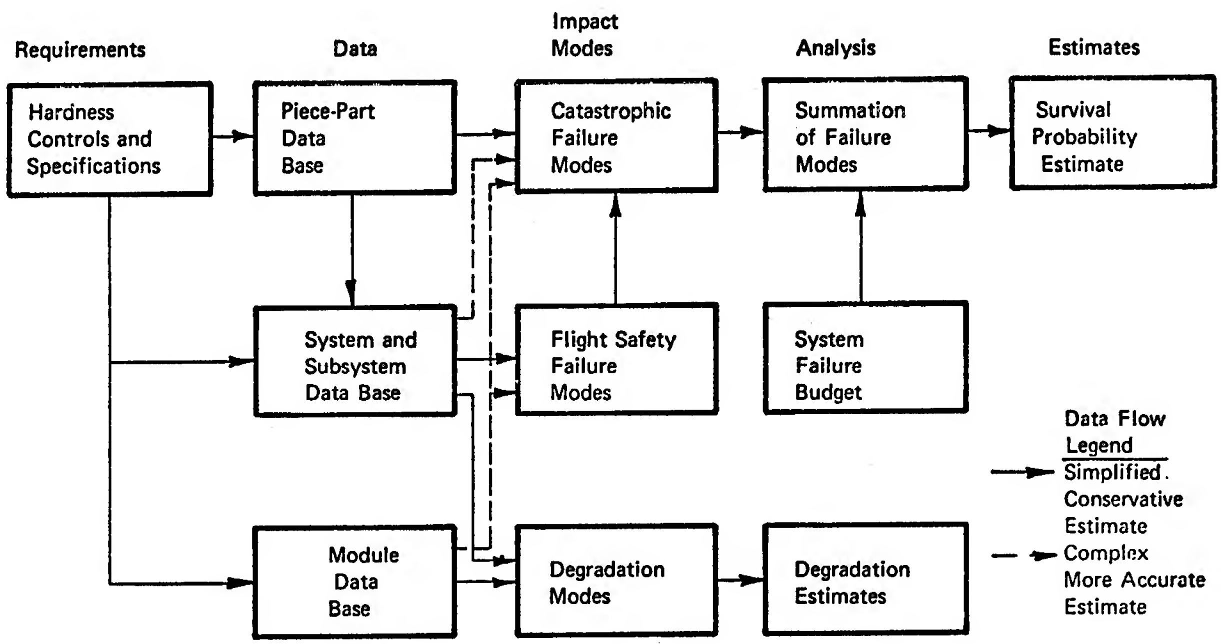


FIGURE 7. INFORMATION FLOW FROM TEST RESULTS THROUGH SYSTEM ANALYSIS (13)

TABLE 1. IMPORTANT VULNERABILITY MODES IN ELECTRONIC SYSTEMS

Failure Mode	Effects
Prompt ionization pulse (gamma and X-rays)	Logic malfunctions Burnout from replacement and ionization current Latchup Recovery time problems Degradation in device characteristics Second breakdown Heating
Delayed ionization pulse (gamma rays, neutrons electrons)	Surface effects - permanent ionization effects Logic malfunction Heating effects Latchup and burnout
Displacement effects (neutrons, gamma rays electrons)	Degradation of semiconductor device performance Lifetime damage Increase in resistivity Rapid annealing effects
EMP (electromagnetic radiation)	Logic malfunction Catastrophic damage - burnout, second breakdown

TABLE 2. TYPICAL TRANSISTOR REQUIREMENTS

Part Number	Operating Current, I_C(mA)	Operating Condition	h_{FE} Min.	V_{ce}(sat) Min.
A3A5	75	Switch	12.2	.5[(a)]
Q4A5	75	Switch	12.2	.5[(a)]
Q1A4	10.3	Switch	8.0	.5[(a)]
Q1A7	900	Switch	8.6	1.7
Q2A4	18	Switch	7.2	.5[(a)]
Q3A4	18	Switch	7.2	.5[(a)]
Q2A8	576	Switch	8.7	1.3
Q3A8	576	Switch	8.7	1.3
Q1A5	20[(b)]	Linear		
Q5A5	(c)	Linear		

(a) Not critical; will cause some increase in h_{FE} requirements if greater than specified.

(b) Emitter current.

(c) The dependence between Q1A5 and Q5A5 does not allow a specific requirement for the current gain of either transistor or the operating current of Q1A5 to be specified. The combined current gain equals 375.

TABLE 3a. ASSEMBLY PROCESSES THAT AFFECT VULNERABILITY AND RELIABILITY (5)

Control Point	Process Controlled or Defect Detected	Advantages Gained
Material		
A. Wire	Wire size and tensile strength	Lead breakage reduction Bond strength
B. Header	Header plating	Bond strength External lead corrosion
C. Die	Die metalization	Bond strength Die attach strength
Die Attach	Preformless	Die failure in special environment
	Process audit	Die attach strength offsets degradation of electrical parameters
Lead Bond		
	Pull test audit	Bond strength
	Bond deformation audit	Bond strength
Preseal Visual	Die attach voids	Die attach strength
	Surface contamination	Intermittence and open and short degradation
	Cracked or chipped die	Die attach strength offsets electrical degradation
	Scratches and metalization defects	Reduces operating life failures
	Oxide and alignment defects	Reduces operating life failures and electrical degradation
	Lead bond deformation and defects	Bond strength

TABLE 3b. ASSEMBLY PROCESSES THAT AFFECT RELIABILITY MORE THAN VULNERABILITY (5)

Control	Failure Mechanisms	Failure Mode
1. Material Control	Metallic inclusions	Intermittent shorts
A. Header	Seal integrity	Hermeticity
B. Cap		Insulation resistance
2. Cleaning-Treating Operations	Surface contamination	Electrical stability
A. Chemical Purity	Foreign material inclusions	Channeling
B. Wash-Rinse Cycles (die & ass'y)		Intermittent shorts
3. Final Seal Operation	Foreign inclusions	Electrical stability
	Internal atmospheric contamination	Opens and shorts
	Internal structural damage	Catastrophic failure
4. Handling-Storage Operations	Surface contamination	Electrical stability
	Internal structural damage	Opens and shorts
5. Assembly Area Control	Contamination	
A. Part Flow (handling-transporting)	Damage	
B. Cleanliness		

TABLE 4. ELECTRICAL MEASUREMENTS USED TO CONTROL TRANSISTOR NEUTRON RESPONSE

Effects	Applicable Measurement
Gain degradation	Base transit time, t_B
	Gain-bandwidth product, f_T
	Emitter-collector delay time[6]
	Preirradiation current gain, h_{FE}
	Emitter depletion capacitance[7]
	Rise-time measurement[8]
	Combination of electrical parameters[9,10]
Saturation voltage increase	Collector breakdown voltage
	Preirradiation current gain, h_{FE}
	Preirradiation saturation voltage, $V_{CE(SAT)}$
	Combination of electrical parameters[9,10]

TABLE 5. CONTROLS AND SCREENS FOR IONIZING RADIATION EFFECTS

Ionization Effect	Control or Screening Procedure
Photocurrent	Baseline process control (mask deterioration, diffusion process, minority carrier lifetime)
	Electrical storage time (t_S)
	Electrical storage time constant (τ_S)[11]
	Inverse gain (β_I)[11]
	Collector depletion capacitance (C_{TC})
	Base spreading resistance, r_B
	Preirradiation testing
Junction burnout	Baseline process control (junction area)
	Passivation control
	Pulse power tests
Metalization burnout	Baseline process control (oxide step control, deposition time, multiple sources, passivation)
	Visual inspection of metalization for voids and scratches
	Visual inspection of metalization with mechanical/electrical aids (e.g., spatial filtering)[12]
	Electrical pulse tests
Latchup	Baseline process control
	Layout control (device design)
	Preirradiation testing
Surface effects	Glassivation
	Stabilization bake (devices and modules)
	Measurement of preirradiation electrical parameters
	Maintain traceability to wafer and/or diffusion run
	Preirradiation testing
	Irradiation/anneal

A SYNERGISTIC RELIABILITY AND MAINTAINABILITY PREDICTION PACKAGE

A.C. Spann

GTE Sylvania Eastern Division
Needham Heights, Massachusetts 02194

INDEX SERIAL NUMBER – 1118

Summary

A mutually compatible set of computer programs for accomplishing routine reliability and maintainability prediction and analysis tasks is a valuable asset. The value realized is proportional to the user's understanding of, and control over, the individual programs and their interaction. This paper describes a particular set of programs in sufficient detail to illustrate the necessary degree of understanding and control. The subject programs are written in the simple BASIC language, and are accessible by keyboard/printer devices via leased time-share service.

Programs ARP and ARPF enable the user to make reliability predictions in accordance with MIL-HDBK-217A. They can also be used in conjunction with a commercially available "canned" program, **RADACF, to make predictions in accordance with the RADC Reliability Notebook. Two primary options are available: (1) parts count by type, individual part stress factor predictions, and (2) total parts count, average stress factor predictions.

Program *AMALA performs two distinct functions: (1) calculation of elemental maintenance task times (fault isolation, remove and replace, etc.) and total mean active corrective maintenance downtime (M_{ct}), based on numerical scores from Checklists A, B, and C of MIL-HDBK-472, Procedure III, and (2) if desired, an analysis of the impact of the projected maintenance concept and environment on the total corrective maintenance downtime.

Also, if desired, *AMALA will create an output file to serve as the input to program *SUMALL. *SUMALL combines the *AMALA results for individual equipments to arrive at projected subsystem or system availability, maintainability, reliability, and logistics parameters. *SUMALL also, upon request, will construct and print a reliability block diagram, with no additional input required.

The cooperative interaction of these programs, with each other and with the user, is illustrated in summary by Figure 1. This stylized flow diagram depicts computer operations as teletypewriter symbols, volatile files as "tailed" circles, and input/output data as sheets of paper. Printed computer output reports are distinguished from user-supplied input data by black corners. Static "stored data" files required by the programs are not shown. A program whose name is preceded by a single asterisk is one developed by, and under absolute control of, GTE Sylvania. The program whose name is preceded by a double asterisk (**RADACF) is commercially available, with user control limited to the options provided.

Introduction

The drudgery of looking up, tabulating, and summarizing reliability and maintainability prediction data has long been the bane of the Reliability/Maintainability Engineer's existence. Now, using time-share terminals and the simple BASIC language most of the drudgery has been eliminated and work can be done exactly as required. One is not constrained to any particular prediction technique or data source. This paper describes the kinds of automated prediction and analysis functions which can be performed, having selected a particular set of techniques and data sources. The usefulness of an integrated set of analytical programs is indeed greater than the sum of its parts.

The most widely used reliability prediction methods, as well as the necessary historical parts data, are documented in: (1) MIL-HDBK-217A, Reliability Stress and Failure Rate Data for Electronic Equipment[1] and (2) RADC TR-67-108, RADC Reliability Notebook[2]. The most authoritative and widely used maintainability prediction procedures are documented in MIL-HDBK-472, Military Standardization Handbook, Maintainability Prediction.[3]

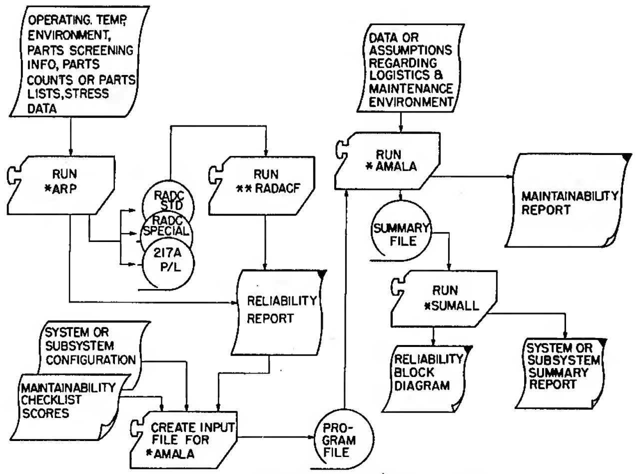


Figure 1. Automated R/M Analysis

The methods and data of these reliability and maintainability handbooks provide adequate but laborious means for making reliability and maintainability predictions for electronic equipment. We will describe here, automated techniques which we have developed for performing these predictions and for analysis of system/subsystem reliability, maintainability, and logistics parameters.

Reliability Prediction

The predicted reliability of electronic hardware is recognized as a useful design tool, being the only available quantitative measure of the expected reliability of hardware under development.

Predicting the reliability of an electronic equipment consists of relating the equipment's electronic part complement to the observed reliability of electronic parts in previous applications. The most widely used prediction methods, as well as the necessary historical parts data, are documented in MIL-Handbook-217A, Reliability Stress and Failure Rate Data for Electronic Equipment, which has been the "Bible" of the defense electronics industry, in the field of reliability, since 1965. Since it is an industry standard, predictions using this handbook data are useful for judging design alternatives, so long as the relative reliability attributed to various part types are reasonably true, regardless of the absolute accuracy of the data.

Program ARP

ARP, a time share program in BASIC language, enables the user to make reliability predictions for electronic equipment in accordance with MIL-HDBK-217A. A gross flow diagram representing the operation of this program is shown in Figure 2. Two primary options are available:

(1) Reliability prediction based on stress factors and part populations - the most detailed and accurate method of the handbook, and the method requiring the most detailed and specific input data.

(2) Reliability prediction based on total parts count and average stresses - a method combining the best features of the gross methods of Section 4 of the handbook, requiring a minimum of input data.

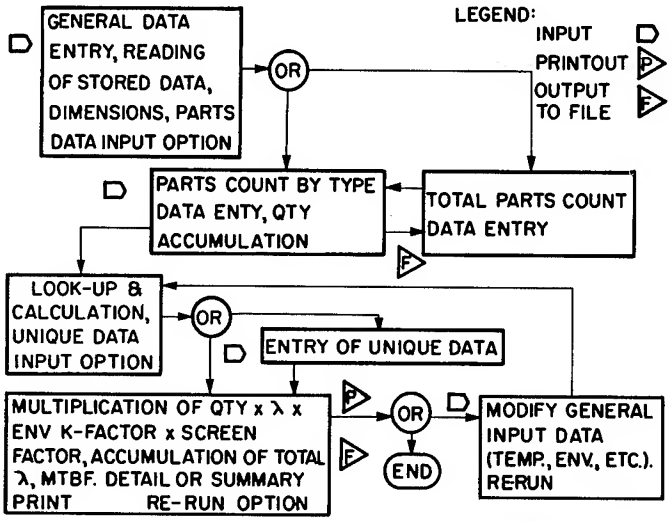


Figure 2. Gross Flow Diagram for Program *ARP

If (1) is chosen, there is an option to enter detailed application and failure rate data for part types not covered by the data in memory (unique parts data). For either primary option, the resulting prediction may be printed out in detail or in summary, at the user's option. (See Figure 4).

Option (1) is intended for making predictions during the design and development of equipments, for use as criteria for detailed design decisions. Option (2) is most useful for application during proposal, pre-proposal, or concept review phases of development, for affecting the gross design decisions (functional configurations, parts de-rating policy, etc.).

Both the detailed and the gross predictions are related to parts population by part type, operating ambient temperature, electrical stresses, and application environment, making maximum use of what is known about these factors in either case.

One aspect not accounted for by MIL-HDBK-217A, i.e., the degree of parts testing and quality control (parts screening), over and above the practices for standard military-grade parts, is provided for in the program.

Procedure In order to perform a valid equipment reliability prediction using this method (or any other method) the equipment assembly level must be chosen, such that all component parts of the assembly are required to operate for successful assembly operation - i.e., the equipment for which the prediction is being made must conform to a series reliability model.

The choice of input option (parts count by type or total parts count input) will depend upon the extent of information available about the equipment. In any case, the average part operating temperature and the intended application environment must be known. The user must also be prepared to state whether, and for what part classes, special parts screening procedures are intended or applied.

Parts Count By Type The Parts Data Format sheet shown as Figure 3 provides a convenient vehicle for assembling a parts count by type. A quantity and average electrical stress (as a ratio of actual to rated) is listed for each part type contained in the equipment. If a given part type has representatives experiencing widely varied stresses, the quantity of that part type may be assigned to as many as six stress groups. Note that the program allows for entry of data for sixty-one specific part types, five "miscellaneous" categories, and up to thirty unique, user-defined part types or categories. Any items in the equipment which are not among the specific part types may be assigned to the appropriate miscellaneous part or assembly category, or defined as unique parts - depending on the available detail of information about these items and the user's desire for accuracy with respect to the failure rates of these items.

Configured Item ____________ Date ____________

Functional Group ____________ Engineer ____________

Component ____________

Average Internal Operating Temperature ___ Estimated / Calculated / Measured

Sequence Number	Part Type	Qty.	Elec. Stress Ratio	Remarks	Sequence Number	Part Type	Qty.	Elec. Stress Ratio	Remarks
1.	Comp. Resistor				39.	XSTR, SI PNP			
2.	Film Resistor				40.	XSTR, SI PNP PWR			
3.	Film Pwr. Resistor				41.	XSTR, SI NPN			
4.	Fixed Acc. WW				42.	XSTR, SI NPN PWR			
5.	Fixed Pwr. WW				43.	XSTR, GE PNP			
6.	G.P. Film Resistor				44.	XSTR, GE PNP PWR			
7.	Var. LD. Scr. Res.				45.	XSTR, GE NPN			
8.	Est. Rel. Film				46.	XSTR, GE NPN PWR			
9.	Var. Comp. Res.				47.	FET			
10.	Var. WW Resistor				48.	UJT			
11.	Var. Non-WW Res.				49.	IC (Digital)			
12.	Paper Capacitor				50.	IC (Analog)			
13.	Mica Capacitor				51.	MOS IC			
14.	But. Mica Capacitor				52.	Transformers			
15.	Var. Cer. Capacitor				53.	Inductors			
16.	Alum. Electro Cap.				54.	Rotary Dev.			
17.	Wet SL. Tant. Cap.				55.	RF Conn.			
18.	SLD. Wet SL. Glass				56.	PC Conn.			
19.	Tub. Fol. Tant.				57.	Rack-Panel Conn.			
20.	Rect. Foil Tant.				58.	Relays			
21.	Solid Tant.				59.	Switches			
22.	Sol. Tant. Est. R				60.	Tubes SM Sig.			
23.	Mylar Cap				61.	Tubes PWR			
24.	Paper-Plas. Cap.				62.	Misc. Parts			
25.	Porce OR GL Cap.					(Lo F.R.)			
26.	Polystyr Cap.				63.	Misc. Parts			
27.	Piston Trmr. Cap.					(Hi F.R.)			
28.	Air Trmr. Cap.				64.	Assembly			
29.	GP Cer. Cap.					(Lo F.R.)			
30.	GP Cer. - T Comp.				65.	Assembly			
31.	Diode SI					(Average F.R.)			
32.	Diode, SI Pwr.				66.	Assembly			
33.	Diode, GE					(Hi F.R.)			
34.	Diode, GE Pwr.								
35.	Varactor					Unique Parts Data Part Name	Qty.	Failure Rate	Environmental K-Factor
36.	Zener								
37.	Microw MXR Diode								
38.	Microw DET Diode								

Figure 3. Parts Data Format

```
FAX FSK CONV        9/18/72              A.C.SPANN

FIXED GROUND ENVIRONMENT

40 °C AMBIENT

  PART TYPE          STRESS   FAILURE RATE    K OP       QTY     TOTAL F RATE
COMP RES               .30    3.50000E-09     6.000       36     7.56001E-07
FILM RES               .30    1.47891E-07     0.030       11     4.88040E-08
   ϟ                    ϟ          ϟ            ϟ          ϟ           ϟ
PC CONN                .20    6.40001E-07     1.100       17     1.19680E-05
RACK-PANL CONN         .20    1.15000E-06     1.100        9     1.13850E-05
SWITCHES               .20    1.10000E-07     1.000        3     3.30000E-07
METER                  ***    5.00000E-07     1.000        1     5.00000E-07
FUSES                  ***    1.00000E-07     1.000        2     2.00000E-07
LAMPS                  ***    1.00000E-06     1.000        3     3.00000E-06

TOTAL                                                    152     1.18845E-04

                    MTBF =  8414.3  HOURS

                      DETAILED PRINTOUT FROM ARP
```

```
                           <<< FAILURE RATES >>>

MODULE # 1          FAX FSK CONV

CIRCUIT
 REF            PART                    FAILURES/MILLION HRS
DESIG           CODE    QUANTITY    UPPER QUAL    LOWER QUAL  DERATING   STRESS

COMP RES        RC07       36        0.00457       0.02115     0.300     0.300
FILM RES        RN60C      11        0.00326       0.01124     0.300     0.300
   ϟ             ϟ          ϟ           ϟ              ϟ          ϟ         ϟ
TRANSFORMER     XFMR        3        0.05400       0.17100     0.300     0.300
INDUCTORS       XL          1        0.04200       0.12000     0.300     0.300
PC CONN         XC2        17        0.02235       0.08940     0.300     0.300
RACK-PANL C     XC3         9        0.03300       0.13200     0.300     0.300
SWITCHES        XSW         3        0.03000       2.25000     0.300     0.300
METER           X01         1        0.50000       0.50000     0.300     0.300
FUSES           X02         2        0.10000       0.10000     0.300     0.300
LAMPS           X03         3        1.00000       1.00000     0.300     0.300

                          <<< EQUIPMENT MTBF >>>

             (UQG)  97251.2     HOURS    (LQG)  25589.1      HOURS
                    DETAILED PRINTOUT FROM **RADACP
```

Figure 4. Reliability Report - Example

Once the user has, (1) decided on the parts count by type option, (2) determined, estimated or assumed a temperature, application environment, and parts screening policy, and (3) completed the parts data format, he is prepared to run the program and answer all requests for input data. After response has been made to all input data queries, the user will be asked if he wishes detail. If the answer is "No", the total parts count, total equipment failure

rate, and equipment Mean Time Between Failures will be the only significant data printed. If the answer is anything other than "No", all significant data for each part type in the equipment will also be printed. When this question is answered, opportunity is given to set the paper to a fresh sheet. Hitting the carriage return will initiate printout.

After the printout, an opportunity is given to RE-RUN. Choosing the re-run results in the opportunity to change the temperature, stress, and parts screening inputs, after which the prediction is performed again using the original part quantities and stresses together with the altered data. Opportunity is also given to change the "unique data" (or to enter unique data, if none were entered in the initial run). If the unique data is to be altered, new data must be entered (or the original data re-entered) beginning with the first line of unique data originally entered, down through the last line to be altered. If the opportunity to alter the unique data is refused, it will be included in the re-run results exactly as it was in the initial run. The same printout option (detail or summary) is offered, and the pause for setting the paper, prior to printout of the re-run results. After printout, the re-run opportunity is again presented. The program will terminate when this question is answered "No"

Total Parts Count For a prediction based on the total parts complement of the equipment, the user need only establish (1) the total parts count, (2) the type of equipment (digital, etc.) and, (3) the over-all average electrical stress on the parts - in addition to the aforementioned temperature, environment, and parts screening policy.

The user is then prepared to answer the self-explanatory requests for input, and the prediction and printout process occurs exactly as described for the parts count by type input.

Under this option, the re-run (if selected) provides opportunity to change all significant input data except the total parts count, then proceeds exactly as before. (Note that unique parts data is not possible with the total parts count option and, therefore, is not involved in the re-run.)

Files ARP requires three files of input data. These files are "static" in the sense that they contain basic data which needs updating infrequently. The program creates three output files. These files are "volatile", since the set of three files are created anew each time ARP is RUN. These output files are named by the user (so that the results of multiple RUNS can be saved, under different file names, if desired). The salient characteristics of the input/output files are as shown in Table 1.

Stored Program Data Since failure rates are on file (in PREDFILE) for each part type, at different stress levels, but at 25°C only, a set of sixty-six failure rate vs. temperature coefficients, K (I), are stored in data statements of the program.

An additional set of data stored in the program is a set of typical fractional part populations. These data are used only for the "total parts count option". They represent the typical fraction of total parts population contributed by part type number one, number two, etc. There are three subsets of these data (sixty-six fractions each) to differentiate between typical parts complements for digital, low-level analog, and high-power equipments.

The part failure rate data of MIL-HDBK-217A apply to standard MIL-grade parts, and do not account for any gains from any pre-conditioning or screening (page 7-1 of MIL-HDBK-217A). It is very desirable to account for the effects of such extra attention to parts quality and reliability, since this is a frequently used method by which manufacturers attain improved equipment reliability. This program accounts for these effects by application of "reliability improvement K-factors" developed by GTE Sylvania[4]. In the event that information is input that no special parts screening procedures are in effect, the standard failure rates are used without modification.

Calculations Given the stored data described in the preceding subsection, and given parts count, stress, and environmental information supplied by the user, the program calculates appropriate part and equipment failure rates. Part type quantities are accumulated to form a total parts quantity. A failure rate is calculated for each part type specified by the input, under the input operating temperature and electrical stress conditions. This calculation is graphically shown in Figure 5. For each part type, the program looks up the two 25°C failure rates which are above and below the stress ratio specified for that part type (for instance, if $s = .4$, the 25°C failure rates at .5 and .3 stress will be selected). Linear interpolation is performed to get the 25°C failure rate at the specified stress ratio, $\lambda_{(25°,S)}$. The part failure rate at the specified temperature, t, and specified stress ratio, s, is found by solving the formula:

$$\lambda_{(t,s)} = \lambda_{(25°,S)} \times 2^{\frac{(t-25)}{K}}$$

where K is the stored coefficient of failure rate vs. temperature.

TABLE 1. CHARACTERISTICS OF INPUT/OUTPUT FILES

FILE NAME OR SYMBOL	REF. NO.	DESCRIPTION	FORMAT
PREDFILE (input)	1	static file of 217A λ at 25°C, env. K-factors, in seq. # order	"PART NAME", $\lambda_{(.1)}, \lambda_{(.3)}, \lambda_{(.5)}, K_{LO}, K_{GF}, K_{AI}, K_S$ ⋮ for 66 part types
RADCODF (input)	3	static file of **RADACF part codes, in ARP seq. # order, + routing indicator to W$	routing indicator*, "**RADACF PART CODE" ⋮ for 100 part types supplementary data unnecessary *1=Λ, 0=W$ data input required
TRANSFL (input)	6	static file of RADC λ's & K-factor for part types of ARP not std to **RADACF	"PART NAME", $\lambda_{(.1)}, \lambda_{(.3)}, \lambda_{(.5)}$ $K_{LO(UQG)}, K_{LO(LQG)}, K_{GF(UQG)}, K_{GF(LQG)}$ $K_{AI(UQG)}, K_{AI(LQG)}, K_{SL(LQG)}, K_{SL(LQG)}$
G$ (input) E$ (output) (TW017A)	7 5	parts data input file (G$) for ARPF created manually or as an ARP output (E$)	"MODULE NAME (P$)" part seq. #, qty, stress ratio ⋮ 70 (signals end of file, unless R$≠"no") R$ (unique parts, "yes" or "no"), A9 "PART NAME", qty, λ, env. K-factor ⋮ for A9 unique part types, if R$ ≠ "no"
V$ (output) (RADSTD)	2	file of standard input parts data for **RADACF	"MODULE", qty (1), "MODULE NAME" "PART NAME", qty," **RADACF PART CODE" or "PT NAME", (-) qty, data per instructions[5] ⋮ "LAST", 0, "0"
W$ (output) (RADSPCL)	4	supplementary file of special parts data for **RADACF	"SPECIAL PT CODE", λ UQG, λ LQG ⋮ "END", 1,1

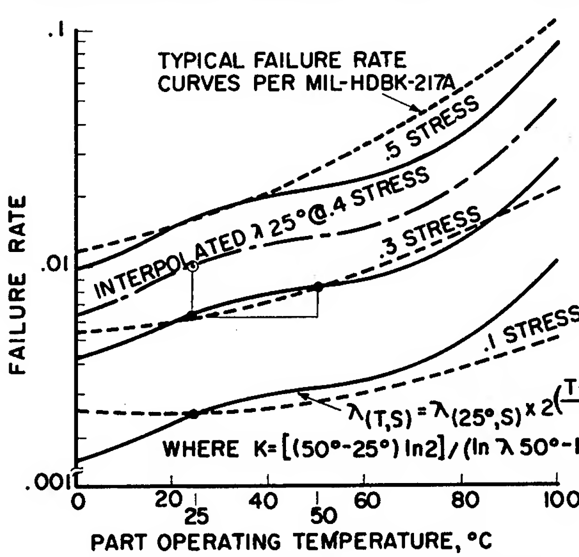


Figure 5. Comparison of Program - Calculated Failure Rates with MIL-HDBK-217A

Once the appropriate failure rate has been established for each part type specified, it is multiplied by the appropriate quantity, environmental K-factor, and if specified, divided by the appropriate parts screening factor to find the total contribution of the part type to the equipment failure rate. These values are accumulated to form the total equipment failure rate, which is inverted to give the equipment Mean Time Between Failures.

Re-Run The re-run is one of two possible operations, depending on whether the initial run was a parts count by type or a total parts count option.

If the initial run was for a parts count by type, opportunity is given to change the specified temperature, environment, and parts screening conditions. Opportunity is given also to make any desired changes to any unique parts data. The prediction process is repeated, using the new data, together with the original part types, quantities and stress ratios.

If the initial run was for a total parts count, opportunity is given to change all significant input data except the total parts count. The prediction process is repeated in exactly the same manner as the initial run, using the new data.

Program ARPF

Program ARPF is identical to program ARP, with the following exceptions:

1. The "total parts count data entry" option is not available.
2. Part type, quantity, and stress data is entered via a pre-prepared file - either the "217A P/L File" created as an output of ARP, or a file in the same format created by any means.
3. No "217A P/L File" is created by ARPF, since such a file must already exist in order to RUN ARPF.

ARPF is considerably quicker and easier to operate than ARP, given that the input data file exists. It is preferred to ARP once initial iterations have been accomplished and the hardware divisions and parts lists "firmed up".

Program **RADACF

**RADACF is a commercially available time-share program which enables the user to perform equipment reliability predictions in accordance with the RADC Reliability Notebook. It requires files of input data describing the equipment parts content and stress data. Companion programs are available, which allow direct entry of this data, much in the manner of ARP. Our interest here, however, is in **RADACF because ARP and ARPF have been designed to translate their input data into output files which are accepted by **RADACF as input files. This eliminates the input preparation process for **RADACF and allows a single, standard input data format for reliability prediction. More importantly, ARP or ARPF processing automatically generates temperature/stress/environment - related failure rates for those part types which must be entered into **RADACF via the "supplementary part code/failure rate file". Otherwise, the **RADACF user must manually create a supplementary file for each temperature/stress/environment combination for which he wants a prediction. It should be noted that ARP enables exercising the "total parts count" option through **RADACF.

Maintainability Prediction

Equipment maintainability predictions serve two primary purposes: (1) design analysis criteria for improving inherent designed-in equipment maintainability characteristics, and (2) in combination with reliability figures of merit, assisting in planning for proper logistics support. The most authoritative and widely-used maintainability prediction procedures are documented in MIL-HDBK-472, "Military Standardization Handbook, Maintainability Prediction". Of the four maintainability prediction procedures described therein, we prefer Procedure III, particularly for application during design and development of systems and equipment. This procedure, in our opinion, is least influenced by subjectivity on the part of the analyst. Further, the resulting maintainability prediction is more easily related back to the mechanical and electrical characteristics of the equipment, making avenues of maintainability design improvement more apparent.

Program *AMALA

Program *AMALA performs two distinct functions: (1) The program will calculate the basic maintenance task times (Localization, isolation, remove and replace, adjust and align, checkout) and total mean active corrective maintenance downtime (M_{ct}) for any specified number of equipments for which input data is provided. The input data consists of numerical scores for Checklists A, B, and C of MIL-HDBK-472, Procedure III. The M_{ct} is calculated in accordance with Procedure III. The basic maintenance task times are calculated by a consistent method which GTE Sylvania developed. (2) If desired, the program will perform an analysis of the impact of the projected maintenance philosophy and environment, and projected logistics support provisions, on the total corrective maintenance downtime per outage (downtime including spares delays, etc., as well as active maintenance time). This analysis is performed using the previously calculated M_{ct}, together with additional maintenance and logistics parameters entered by the user. The object of this routine is to enable the user to examine the projected maintenance and logistics environment, together with the equipment maintainability characteristics, to see if it is satisfactory. If it is not satisfactory, he can re-iterate with changing conditions until he finds a satisfactory combination.

As an aid to this iterative process, a problem diagnosis routine will aid in isolating the problem (excessive projected downtime) to either the equipment maintainability characteristics or the maintenance and logistics environment. This routine will delve further to isolate to a particular design aspect or logistics provision, as the case may be. This provides guidance for the user in re-running to seek a satisfactory situation.

If desired, the program will create an output file to serve as the input to program *SUMALL.

An over-all description of program *AMALA is provided by Figure 6.

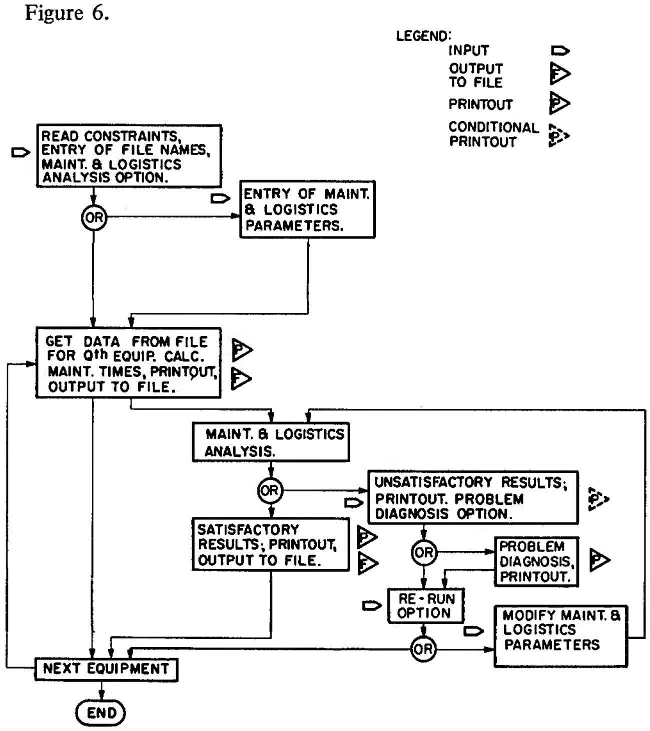


Figure 6. Gross Flow Diagram for Program *AMALA

Procedure In order to run *AMALA, an input data file must be created. The format for this file is as follows:

Number of equipment types in file
Equipment name
Failure rate, Qty., Qty req'd for success
Checklist A scores (15 items)
Checklist B scores (7 times)
Checklist C scores (10 items)
Equipment name

.
.
.

etcetera

Checklists which must be scored to provide this data are included in Appendix A of MIL-HDBK-472. If it is intended to run the maintenance and logistics analysis, as well as task time predictions, and to run *SUMALL, an output file name must be created. One can run both the prediction and M & L analysis without an output file by entering "none" as the output file name in response to the *AMALA input request. Refer to Figure 7 for a sample run of *AMALA.

```
FILE NAMES (INPUT,OUTPUT)? CRFACFL,SUMFL
DETAILED PRINTOUT? YES
DO YOU WISH LOGISTICS ANALYSIS, AS WELL AS TASK TIMES? YES
ENTER LOGISTICS DELAY TIMES:
   LOCAL SPARES ACQUISITION (MINUTES)? 10
   I/L REPAIR TIME (MINUTES)? 240
   DEPOT TURN-AROUND (HOURS)? 72
ENTER RANGE OF SPARES AVAILABILITY (MIN,MAX)? .5,.9
MAX. ALLOWABLE MAINT. DOWNTIME (MINUTES)? 30

FSK CONVERTER       TOTAL DOWNTIME ( 30.6 MINUTES)
EXCEEDS THE DESIRED MAXIMUM FOR ALL CONDITIONS.

PROBLEM DIAGNOSIS? YES
***EFFECTS OF THE MAINTENANCE ENVIRONMENT***

% CONTRIBUTIONS TO TOTAL DOWNTIME ARE:

UNAVAILABILITY OF SPARES AND DEPOT TURN-AROUND     28.3%
UNAVAILABILITY OF SPARES AND I/L REPAIR TIME        4.7%
ACTIVE MAINTENANCE + LOCAL SPARES DELAY            67.0%

***EFFECTS OF DESIGN CHARACTERISTICS***

REVIEW CHECKLIST C, ITEMS:
 4
 6
 7
 8
 9
```

AMALA
INTERACTIVE MAINTAINABILITY ANALYSIS

```
FILE NAMES (INPUT,OUTPUT)? CRFACFL,SUMFL
DETAILED PRINTOUT? YES
DO YOU WISH LOGISTICS ANALYSIS, AS WELL AS TASK TIMES? YES
ENTER LOGISTICS DELAY TIMES:
   LOCAL SPARES ACQUISITION (MINUTES)? 5
   I/L REPAIR TIME (MINUTES)? 240
   DEPOT TURN-AROUND (HOURS)? 72
ENTER RANGE OF SPARES AVAILABILITY (MIN,MAX)? .5,.9
MAX. ALLOWABLE MAINT. DOWNTIME (MINUTES)? 35
```

ITEM NAME	MAINTENANCE TASK TIMES (MINUTES) LOC	ISOL	R/R	ADJ	C/O	TOT	SPARES PROB	UNSCHED DNTIME
FSK CONVERTER	0.9	9.3	3.7	0.2	0.9	14.9	.80	34.5
HF RECEIVER	0.7	9.2	3.7	4.1	0.7	18.5	.84	34.9
HF MULTICOUPLER	0.5	6.9	4.2	1.0	0.5	13.1	.78	34.3
HF ANTENNA	0.5	5.4	4.3	0.2	0.5	11.0	.75	34.5

AMALA FINAL OUTPUT REPORT

Figure 7. Maintainability Analysis and Report - Example

The input requests are self-explanatory. It should be pointed out, however, that if it is desired to run *SUMALL:

(a) A previously-named file must exist for use as an output file by *AMALA, and as an input file by *SUMALL.

(b) The *AMALA option to perform maintenance and logistics analysis, as well as task time predictions, must be answered "yes."

(c) The *AMALA analysis must be re-run, if necessary, until a satisfactory situation is achieved for each equipment in the file.

The program has two basic operations, the maintainability prediction routine and the maintenance and logistics analysis option, which contains two sub-options - problem analysis and re-run.

Maintainability Prediction A given item, or question, of the checklists relates to a specific maintenance task or tasks.

Now, let us consider a world in which equipment maintainability is characterized by these checklist scores as they vary uniformly - that is, scores for all items are the same; all 4, or all 3, etc. Scanning the checklists we find that 10 of the 32 scores appear to affect all maintenance tasks. We will set these aside, since they contribute to all tasks, and we are concerned now with the apportionment of total M_{ct} to specific tasks. Of the remaining 22 scores, 4 (or 4/22 of the total score) affect Localization and Checkout, 9 (or 9/22 of the total score) affect Fault Isolation, 8 (or 8/22 of the total score) affect Remove and Replace actions, and 1 (or 1/22 of the total score) affect Adjustment.

The formula for total corrective maintenance downtime (M_{ct}) is:

$$M_{ct} = \text{antilog}_{10}\,(3.54651-.02512A-.03055B-.01093C)$$

Where A is the total score for Checklist A
B is the total score for Checklist B
C is the total score for Checklist C

It seems reasonable then that the components of M_{ct} could be expressed in the same form;

$$\text{Task Time (e.g., Isolation)} = \text{antilog}_{10}\,(E\text{-}XA\text{-}YB\text{-}ZC)$$

Where A is the sum of scores for Checklist A items affecting Isolation time.
B is the sum of scores for Checklist B items affecting Isolation time.
C is the sum of scores for Checklist C items affecting Isolation time.
E, X, Y, and Z are constants.

Having solved for the necessary constrants,[6] in addition to the given relationship for M_{ct}, we can say that:

$$T_{LOC} = T_{C/O} = \text{antilog}_{10}\,(2.50512-.261992A_1-.318625B_1)$$

$$T_{ISOL} = \text{antilog}_{10}\,(3.15833-.059372A_2 -.072207B_2-.025834C_2)$$

$$T_{R/R} = \text{antilog}_{10}\,(3.10718-.072452A_3 -.088114B_3-.031525C_3)$$

$$T_{ADJ} = \text{antilog}_{10}\,(2.20409-.653548A_4)$$

Where A_1, B_1, etc. consist of the sums of the appropriate checklist item scores as defined earlier.

Maintenance and Logistics Analysis The maintenance and logistics analysis is performed in terms of a downtime model depicted by Figure 8. The elemental task times t_1, t_2, t_5, t_9, and t_{10} are calculated by the prediction routine described in the preceding section. Other time elements must be supplied by the user. The boxes of Figure 8 represent time-consuming maintenance tasks or logistic elements; the ovals represent decision points having outcomes which can be stated probabalistically.

Note that there are ten possible paths leading from malfunction through checkout. We will designate these paths as:

Path A composed of time elements t_1, t_2, t_3, and $t_{10} = t_A$,
Path B composed of time elements t_1, t_2, t_3, t_9, and $t_{10} = t_B$,
Path C composed of time elements t_1, t_2, t_5, t_8, and $t_{10} = t_C$,
etc.

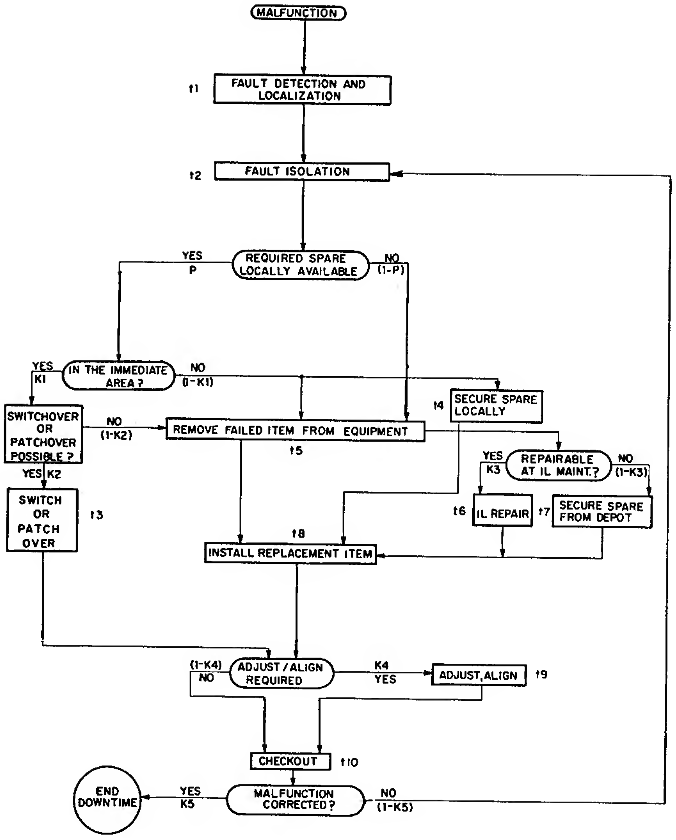


Figure 8. Maintenance and Logistics Analysis

Notice that t_1, t_2, and t_{10} are common to all ten paths - i.e., there is no way to avoid these three elements of time. They will be present for all corrective maintenance actions. The design must therefore provide for minimization of the time required for these three tasks.

The possible maintenance paths are listed in order of preference. The relative frequency with which a preferred path is available is determined by the "P" and "K" probability factors associated with the decision outcomes.

These factors, in turn, are determined by the equipment design and the logistics support environment provided.

These probability factors are defined as follows:

P is the probability that a spare item is available when required, regardless of the failed item type.

K_1 is the probability that a spare item is stored in close proximity to the failed item.

K_2 is the probability that a spare item can be immediately switched or connected into the equipment to restore operation, without first removing the failed item.

K_3 is the probability that a failed item is repairable at Intermediate Level maintenance.

K_4 is the probability that the replacement of a failed item will cause adjustment or alignment of the item or other items of the equipment to be required before returning to service.

K_5 is the probability that a completed corrective maintenance action will be successful in restoring the equipment to operation.

The probability that a given path will be taken is given by the cumulative product of all the probabilities found along that path. Therefore:

$$P_A = PK_1K_2(1-K_4) \qquad P_F = P(1-K_1)K_4$$
$$P_B = PK_1K_2K_4 \qquad P_G = (1-P)K_3(1-K_4)$$
$$P_C = PK_1(1-K_2)(1-K_4) \qquad P_H = (1-P)K_3K_4$$
$$P_D = PK_1(1-K_2)K_4 \qquad P_I = (1-P)(1-K_3)(1-K_4)$$
$$P_E = P(1-K_1)(1-K_4) \qquad P_J = (1-P)(1-K_3)K_4$$

Note that, as expected, $\Sigma P = 1.0$, which is to say – if you enter maintenance (given a malfunction), you will emerge.

At this point, we can express the long-term mean total unscheduled downtime as:

$$D_t = P_At_A + P_Bt_B + \ldots + P_Jt_J = \sum_{i=A}^{J} P_it_i$$

Problem Analysis In the event that the maintenance & logistics analysis results in an unsatisfactory situation (downtime greater than the specified allowable maximum), the option to perform problem diagnosis will be offered. If answered "yes," this routine will check to see if the active maintenance time (M_{ct}) exceeds the specified maximum total downtime.

(a) If not, the program will calculate and print out the percentage contribution to total downtime of all the maintenance paths of Figure 8 and then examine the checklist scores, printing out those checklist item numbers which contribute most heavily to poor design maintainability.

(b) If so, the program will not examine the maintenance path (delay) contributions, but will proceed directly to the maintainability design analysis.

If the problem analysis option is answered "no," the opportunity to re-run will be offered.

Re-Run If the re-run option is answered "yes," opportunity is given to change any or all maintenance and logistics parameters previously entered by the user. The maintenance and logistics analysis will be repeated exactly as described previously, using the modified data and the original maintenance task times. Problem analysis and re-run options will be offered in each re-run until a satisfactory situation (predicted total downtime less than or equal to the specified maximum allowable total downtime) is achieved.

System/Subsystem Summary Analysis

The previously described parlay of inter-active programs enables the user to assess the reliability and maintainability of arbitrary groupings of electronic parts. Typically, one would choose the line replaceable unit (LRU) as the level for iterative analysis and prediction of reliability using ARP, ARPF and **RACACF, and assemble the resulting reliability predictions, along with the checklist scores and system or subsystem configuration information into the input file for analysis via *AMALA. Program *AMALA provides maintainability characteristics for the individual LRU's, recognizing them as members of the system or subsystem only to the extent reflected by the maintainability checklist scores. However, the output file created by *AMALA contains all the data required for a comprehensive description of the reliability, maintainability, and logistics characteristics of the system or subsystem.

Program *SUMALL

Output reports describing these system/subsystem characteristics are generated by program *SUMALL. Two reports are available. The first is a tabular listing of system components (black boxes) and their associated quantities, failure rates, and mean time

to repair, followed by a series of statements of system reliability, maintainability and availability characteristics - and projected logistics support requirements (see Figure 9). The second, optional report consists of a completely annotated reliability block diagram (Figure 10).

```
RADIO FACSIMILE

     COMPONENT           QTY    FAILURE RATE     MTTR
                               (FAILURES/HR.)  (MINUTES)

FSK CONVERTER             1    1.18845E-04      14.9
HF RECEIVER               1 *  1.61455E-07      18.5
HF MULTICOUPLER           1    2.17000E-05      13.1
HF ANTENNA                1    8.30000E-06      11.0

TOTAL FAILURE RATE             1.49006E-04

THE MEAN TIME TO REPAIR IS  14.4 MINUTES

MEAN TIME BETWEEN FAILURES IS    6711 HRS.

THE INHERENT AVAILABILITY IS  .999964

THE MEAN DOWNTIME PER OUTAGE IS  34.5 MINUTES,
GIVEN A MINIMUM SPARES AVAILABILITY OF  .79
APPORTIONED AMONG THE COMPONENTS IN ACCORDANCE
WITH THE RESULTS OF THE LOGISTICS ANALYSIS.

THE OPERATIONAL AVAILABILITY IS  .999915

MEAN TIME BETWEEN MAINTENANCE IS    1259 HRS.

ORGANIZATIONAL MAINTENANCE MAN-LOADING IS
 0.5 MAN-HOURS PER 1000 OPERATING HOURS.

      * REDUNDANCY - SEE TEXT.
```

Figure 9.System R/M Summary Report - Example

```
                  |
       -----------------------| MTTR= 14.9 MINUTES
      |FSK CONVERTER          |
      |   QTY OF  1           | AT     118.84500 FAILURES/MILLION HRS
      |_______________________| FAILURE RATE= 118.845     FAILURES/MILLION HRS
                  |
                  |
            ------------ - - -|
           |                  |
 -------------------    | |----------------------| MTTR= 18.5 MINUTES
|HF RECEIVER        |   | |HF RECEIVER           |
|                   |   | |                      | 2     OUT OF 3     REQUIRED
|___________________|   | |______________________| AT 215.     FAIL/MILLION HRS
           |                  | EQUIVALENT
           |__________ - - ___| FAILURE RATE= .161455     FAIL/MILLION HRS
                  |
                  |
       -----------------------| MTTR= 13.1 MINUTES
      |HF MULTICOUPLER        |
      |   QTY OF  1           | AT      21.70000 FAILURES/MILLION HRS
      |_______________________| FAILURE RATE= 21.7     FAILURES/MILLION HRS
                  |
                  |
       -----------------------| MTTR= 11.0 MINUTES
      |HF ANTENNA             |
      |   QTY OF  1           | AT       8.30000 FAILURES/MILLION HRS
      |_______________________| FAILURE RATE= 8.3   FAILURES/MILLION HRS
                  |
                  |
OVER-ALL MTBF=  6711.12     HRS  MTTR=  14.4049      MINUTES

****RELIABILITY BLOCK DIAGRAM - RADIO FACSIMILE****
```

Figure 10. Reliability Block Diagram

__Procedure__ The only prerequisites for running *SUMALL are (1) the name of the input file (*AMALA output file, or a file in the same format) (2) a knowledge of whether non-interfering on-line repair of the system is permitted in context with redundancies (if any), and (3) some arbitrary system or subsystem name. When these questions are answered, opportunity is given to set the paper to a new sheet, and the first report is printed out. The user is then asked if he wishes a reliability block diagram. If the answer is "yes," the paper is again reset and the block diagram is printed. If "no," the program terminates.

__Calculations__ Program *SUMALL makes the following calculations: for each type component (black box) in the system;

Serial failure rate = $n\lambda$

Maintenance man-hrs/1000 operating hrs $= n\lambda N\, M_{dt} \times 1000$

Then, sensing whether $n = n_1$, and whether redundancies are with or without repair;

if $n = n_1$, failure rate = $n\lambda$

if $n \neq n_1$ and redundancy is without repair, failure rate $= n(n-1)\ \lambda/(2n-1)$

if $n \neq n_1$ and redundancy is with repair, failure rate $= n(n-1)\,\lambda^2 / ((2n-1)\,\lambda + (1/M_{ct}))$

Active maintenance downtime/operating hour = M_{ct} X failure rate

Total maintenance down/time operating hour = M_{dt} X failure rate

where n = the quantity of the given component type per system

n_1 = the quantity of that component type required for successful system operation

λ = the per-component failure rate for that component type

M_{ct} = mean active corrective maintenance downtime for the given component type

M_{dt} = mean total corrective maintenance downtime for the given component type

N = the average number of maintenance men per maintenance action, required for the given component

all of which are contained in the input file.

The program then makes the system-level calculations:

Total failure rate = Σ failure rate

$$\text{System } M_{ct} = \frac{\Sigma \text{ Active maintenance downtime/operating hour}}{\text{Total failure rate}}$$

System MTBF = 1/Total failure rate

$$\text{System inherent availability} = \frac{\text{System MTBF}}{\text{System MTBF} + \text{System } M_{ct}}$$

System mean downtime per outage

$$= \frac{\Sigma \text{ Total maintenance downtime/op. hr.}}{\text{Total failure rate}}$$

$$\text{Minimum req'd spares availability} = \frac{\sum_{i=1}^{K} P_i n_i \lambda_i}{\text{Total failure rate}}$$

for the k component types of the system, where P_i is the probability of not running out of spares required for the i^{th} component type, in order that its mean total downtime per outage not exceed M_{dt_i}. (P_i is also contained in the input file.)

System operational availability

$$= \frac{\text{System MTBF}}{\text{System MTBF + System mean downtime/outage}}$$

System mean time between maintenance = $1/\Sigma$ Serial failure rate

Organizational maintenance man-loading per 1000 operating hours $= \Sigma$ Maintenance man-hrs/1000 operating hours.

Note that *SUMALL, for redundancy calculation purposes, always assumes that n-1 of the n components affected are required for successful system operation, regardless of the value of n_1 originally entered by the user in the input file to *AMALA. This is the most practical arrangement in most cases, and any divergence from the truth is in the pessimistic direction.

Conclusions

The application of a set of programs such as this allows for vastly increased effectiveness, as well as efficiency, on the part of the Reliability/Maintainability Engineer. It relieves him of time-consuming, mind-dulling tasks - making additional time and idea-stimulating data available to him. The outputs of the programs are

directly usable in reports, arithmetic errors do not occur, and typographical errors are extremely rare.

Effective as it is, the set of programs described herein is merely a first step toward even more exciting possibilities. Such things as basic data files that "learn" and built-in Bayesian generation of failure rate data for new device types are feasible now. Files that "learn" for instance - recall that ARP contains stored data on typical part type mixes, which is used to synthesize equipment parts counts by type under the "total parts count" option. Suppose we arrange the program so that, when a user is entering what he believes to be a "typical equipment", under the "parts count by type" input option, he triggers the program to output this parts count into an output file. Such outputs could be used to automatically update the "typical parts mix" data, when certain preset conditions are reached.

I have pointed out all significant shortcomings and assumptions inherent in the programs, as they now exist. This is important. Because there is a tendency to grant "instant credibility" to computer printouts, the importance of knowledgeable, responsible use of such programs is magnified.

References

(1) MIL-HDBK-217A, "Reliability Stress and Failure Rate Data", 1 December 1965 (and Notices 1 and 2 thereto, dated 3 November and 30 December, 1971, respectively).

(2) RADC-TR-67-108, "RADC Reliability Notebook, Vol. II", 5 January 1968.

(3) MIL-HDBK-472, "Maintainability Prediction", 24 May 1966.

(4) GTE Sylvania report ACS:1209-71, "Automated Reliability Prediction Program", 23 August 1971.

(5) Service Bureau Corporation publication SJ066, "CALL/360: **RADAC, **RADACF, **RADAC1 Reliability Calculations", 1971.

(6) GTE Sylvania report ACS:1226-71, "Automated Maintainability and Logistics Support Analysis", 30 August 1971.

FORTRAN VS. GPSS IN RELIABILITY SIMULATION

INDEX SERIAL NUMBER — 1119

by

Orlino C. Baldonado
Holmes & Narver, Inc.
Anaheim, California

and

J. Wes Pierson
Litton Data Systems Division
Van Nuys, California

Abstract

The General Purpose Simulation System (GPSS) language has not been used as extensively as Fortran in reliability simulation studies. GPSS has been used mainly for studying the discrete flow of transactions through a system, hence its use in reliability analysis has been indirect. However, it can be used in a direct way. A study was undertaken to compare the relative merits of the two languages for system safety and reliability evaluation.

A fault tree was used to represent the system. In Fortran, the fault tree was inputted as a subroutine after translating it into logical expressions. A component in the failed state would have its corresponding logical variable set equal to 1 or "true"; otherwise it would retain the value of 0 or "false". In GPSS, the components were represented by logic switches, which were set or reset according to whether the component was in the failed state or not. Boolean variables were used to combine the logic switches to represent the fault tree.

Several systems represented by fault trees were simulated. Simulation consisted of randomly failing and repairing the components by the use of time-to-failure and time-to-repair distribution functions. This was accomplished by generating events corresponding to failure or repair and arranging these events to occur in simulated time. The system was checked to determine if it had failed by evaluating the logic subroutine in Fortran and by evaluating the Boolean variables in GPSS.

It is seen that GPSS may be superior to Fortran in some respects, especially when dealing with a few number of components (20 or less). The ability of GPSS to easily handle distribution functions in tabular form, the built-in features for gathering statistics, and the way events are controlled to occur in simulated time are seen as major advantages. Fortran, however, by the use of arrays can handle a relatively larger number of components, and can also have more accuracy in calculating failure probabilities.

It is recommended that the use of GPSS be explored in reliability analysis in addition to the use of Fortran since it has certain advantages that can relieve the analyst of much tedious programming work.

Introduction

There exist many ways of analyzing system reliability. These include failure modes effect and criticality analysis, fault tree analysis, series-parallel block diagram analysis, and system simulation. Some of these methods reflect the way in which the system is represented; others reflect the analytical or computer methods that are applied to the system representation.

Simulation has been an effective tool for reliability analysis. It has been applied to simulate the random failures associated with components in a system represented by fault trees or block diagrams. Fortran has been the main language used. One would expect the predominant use of simulation languages like GPSS, SIMSCRIPT, GASP, SIMULA, etc., but such is not the case. In particular, the general purpose simulation system (GPSS), although it is a widely used simulation language, is not a natural language for

direct reliability simulation since there is no readily identifiable transaction that flows through the system. In fact, the author is not aware of any direct use of GPSS in system reliability simulation. However, GPSS has been used extensively in the evaluation of system performance, including the effects of failures on this performance, and therefore, its use in reliability analysis has been indirect.

This paper will discuss some experience related to the use of GPSS in some applications where Fortran has been the dominant computer language used. It will be seen that GPSS can be superior to Fortran in certain respects.

System Representation by Fault Trees

Consider a system and associated with it is a system failure that one can identify. A fault tree is often used to display the inter-relationship of fault events leading to the system failure. The system failure is called the undesired event. A fault tree is a diagram of fault events leading to system failure. It is a graph which delineates all components or events and their relationship to the undesired event.

In constructing a fault tree, the undesired event is called the "top" event. The subevents that lead to the top event are identified. The subevents are further traced to sub-subevents that lead to them. The result is a graphical representation of the possible sequences of events that lead to the top event. Those events that lie at the end of the fault tree are the basic input events. In most cases, these input events correspond to the failure of components with an identifiable meantime to repair and repair time.

To illustrate what is being discussed, a "two-out-of-three" system is used as an example. This is a system where there are three components, at least two of which are required to keep the system in operation. The system fault tree for such a system is shown in Figure 1.

A gate indicated by V is an OR gate; one indicated by Λ is an AND gate. The circles with numbers indicate component states. A fault tree for a larger system contains other types of gates and other symbols that make it possible to display event relationships.

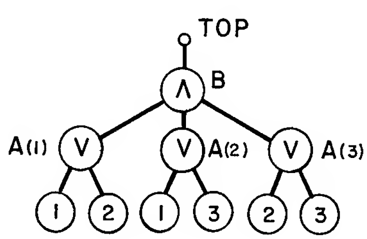


FIGURE 1

Two-out-of-Three System Fault Tree

Fortran System Representation

In Fortran, logical variables are associated with the top event, the gates, and the input events. A value of 1 or "true" corresponds to the failed state and a value of 0 or "false" corresponds to the non-failed state.

A set of logical statements relating to the top event were made into a subroutine (called subroutine logic) to represent the system. It is necessary to change this subroutine everytime there is a different fault tree corresponding to a different system.

The subroutine for the example two-out-of-three system is shown in Figure 2.

```
SUBROUTINE LOGIC (X,TOP)
LOGICAL A, X, TOP
DIMENSION A(3), X(3)
A (1) = X(1) .AND. X(2)
A (2) = X(1) .AND. X(3)
A (3) = X(2) .AND. X(3)
TOP = A(1) .OR. A(2) .OR. A(3)
RETURN
END
```

FIGURE 2

Subroutine Logic for a Two-out-of-Three System

GPSS System Representation

Logic switches take the place of components in a GPSS simulator. A logic switch is set or reset corresponding to the failure or non-failure of the components. Boolean variables are used to combine the logic switches similar to that in Fortran. The set of logic switches and Boolean variables were incorporated as statements into the simulation model. It would be necessary to change these statements whenever there is a different faulttree that is being simulated.

The GPSS blocks for the example two-out-of-three system are shown in Figure 3.

```
1      BVARIABLE LS1*LS2
2      BVARIABLE LS2*LS3
3      BVARIABLE LS1*LS3
4      BVARIABLE BV1+BV2+BV3
```

FIGURE 3

GPSS Boolean Variables for a Two-out-of-Three System

Fortran Fault Tree Time Simulation

Initially, all inputs to the fault tree are set equal to false, i.e., all components are assumed to be operating. Random numbers corresponding to the mean-time-to failure of the components are generated using the component failure distribution functions, and these numbers are placed in an array. The elements of this array initially correspond to the first time at which the components fail. The Fortran program sorts

these elements starting from the smallest value to the largest value. It then takes these sorted elements, and steps time from zero to the smallest value in the array, and sets the value to true of the logical variable corresponding to the input with the smallest random time to failure. This corresponds to failing of one of the components, i.e., a failure event takes place. Subroutine Logic is next examined to determine if the system has failed or not. If it has failed, (i.e., the top event is found to be true) the single failure event is noted as causing system failure and the simulator stores the value of time and proceeds.

A random time-to-repair for the failed component is next determined, using the component repair distribution function. The component is assumed to be in the failed state from the time it has failed to time plus the random time to repair. Subsequent examination of Subroutine Logic during this time interval will show that this component is in the failed state. This is accomplished by placing the time at which a failed component is repaired into the array, resorting the elements of the array, and setting the logical variable to false only at this time.

Meanwhile, other components are being failed and repaired. Any failure or any repair time is an event. The time of occurrence of these events are placed in the array. The simulator sorts all these events in time, and goes on to examine the state of the system everytime there is an event.

Times at which the system is failed and the system is repaired are stored for further processing. There is a certain system real time limit that is reached, at which time the simulator starts all over again, resetting time to zero with all components in the unfailed state. Many trials are also performed.

A tabulation of the system time to failure and system time to repair gives the desired information on the system behavior. Single component failures, or combinations of components failing and leading to the system failure (the undesired event) are also saved and printed out to determine which components are more critical than others.

GPSS Fault Tree Time Simulation

Initially, all logic switches in the simulator are reset. This corresponds to all components operating. Also, save values are set up which contain the mean times between failures (MTBF) and the mean times to repair (MTTR) for each component.

At the start of the simulation, a transaction is created for each of the components in the modeled system. These transactions enter an ADVANCE block which uses the simulator's random number generator and the inputted MTBF to determine the time at which this particular component will fail. When this point in simulated time (which is different for each transaction) is reached, the transaction leaves the ADVANCE block and the component's corresponding logic switch is set. This puts the component in the failed state. If the total system is in the operating state, Boolean variables are evaluated to determine if this component will cause the system to fail; otherwise, this action is skipped. The transaction then enters another ADVANCE block which calculates the component's MTTR. When the transaction leaves this block, the corresponding logic switch is reset putting the component back in the operating state. Then if the total system is in the failed state, the Boolean variables are re-evaluated to determine if the system has been repaired. The transaction then starts the process over again.

The GPSS simulator automatically keeps track of all events as they occur in time and tabulates MTBF's and MTTR's distribution functions for the total system as directed by TABULATE blocks.

A simulation timer is used to control how much simulated time is to elapse before the end of the run. At this time, new input values may be read in, changing one or more of the component MTBF's and/or MTTR's and the simulation started over.

Results and Findings

The following summarizes the findings on the relative merits of the two languages with respect to this type of application:

(1) For this type of reliability simulation, GPSS was more simple to program. The Fortran programming effort is considerably less than that of the GPSS effort. In fact, the number of Fortran statements is about five times that of the GPSS statements. The authors are equally proficient in both Fortran and GPSS.

(2) Once the programs are set up, it is relatively easy to update both the Fortran program and the GPSS program.

(3) The input data is part of the GPSS program (by the use of initial blocks). In Fortran, they are read in as data cards. A change in data requires reinterpretation of the GPSS program which can be a disadvantage.

(4) For systems with few numbers of components (20 or less) it is easier to structure the model in GPSS than in Fortran. For a system fault tree with more than one hundred input events, Fortran has a definite advantage.

(5) The ordering of the events is automatically performed in GPSS (and also more efficiently). It takes a large part of the computational time in Fortran. GPSS has the advantage over Fortran in this respect.

(6) Tabulations of events are easier to perform in GPSS than in Fortran.

(7) By use of some of the computational and arithmetic features in Fortran and not available in GPSS, Fortran gives more accuracy as far as system failure probability is concerned.

(8) Fortran is more accessible than GPSS, independent of the application being considered.

(9) The GPSS model runs faster than the Fortran models.

In general, it is the presence of built-in features of statistical tabulation, and event-ordering that gives GPSS some advantages over Fortran.

Conclusions

Simulation will continue to be a useful tool in analyzing system reliability via fault tree analysis. The use of specialized languages like GPSS instead of just Fortran can have advantages. It is recommended that its use be explored by the system reliability analyst.

MODEL TO MAKE ARMY DECISIONS

LEWIS NERI
Chief, Reliability and Maintainability Div
Directorate for Product Assurance
US Army Aviation Systems Command
St. Louis, Missouri

DR. HENRY WIEBE
Associate Professor
Engineering Management Dept.
University of Missouri-Rolla
Rolla, Missouri

INTRODUCTION AND SUMMARY

The Engineering management decision technique currently in use by the Army's Aviation Project Engineers to determine which Equipment Improvement Recommendation or EIR case should be evaluated first, has been studied and a computer program designed to perform this function. Four significant parameters - reliability, availability, total annual inventory cost and total annual cost to live with the problem - have been developed and used to accomplish this. The objective of this study was to computerize the manual and mental process and evaluation of the EIRs relative to the four parameters and arrive at the decision as to which EIR has the highest priority.

BACKGROUND AND CURRENT DECISION PROCESS

The US Army Aviation Systems Command (AVSCOM), located at 12th and Spruce Streets, St. Louis, Missouri, is a major subordinate command reporting directly to the Army Materiel Command (AMC) which in turn reports to the Department of the Army (DA). The responsibility of AVSCOM is total management for assigned aviation systems and items, including all interfaces with other commodity commands.[1] Specifically, it develops and provides worldwide aviation materiel and related technical, professional guidance and assistance required for the support of the Department of the Army Aviation Materiel and other U. S. and foreign customers. Once the materiel is procured or assigned, AVSCOM plans and conducts new equipment training, special training, including the training of foreign nationals.

In order to carry out its support mission, AVSCOM establishes systems project offices with an Army Project Engineer for each aircraft system. The mission of the aircraft project offices are to provide the engineering required to assure the integrity and reliability of fielded Army aircraft and ground support equipment, armor systems, materials, avionics and other installed systems.[2]

One of the ways which the Army Project Engineer uses to determine the problem areas and accomplish its assigned mission is through the Army Integrated Equipment Record Maintenance Management System commonly referred to in the Army as the TAERS System. The data feedback system currently is not completely satisfactory but it is acceptable. An area for improvement exists due to the fact that better utilization could be made of the incoming data. The data collected under the TAERS system was not utilized to its utmost because management techniques for the effective utilization of the Army's Aviation Rotary-Wing Reliability and Maintainability and In-House Data Collection Programs had not been adequately developed and thus have not occupied a prominent place in the work of the Army Project Engineer. There have been two major reasons for their reluctance to engage in this endeavor. First, the apparently overwhelming accumulation of myriad data, and formidable and tedious tasks involved in understanding the stochastic techniques used for analysis. Second, the Army Project Engineer did not have the time or resources to develop these techniques as he was responsible for providing engineering support to the fielded aircraft systems, providing contractual technical requirements, evaluating equipment improvement recommendations (EIR), preparing technical studies, developing and evaluating both in-house and contractor Engineering Change Proposals (ECP) developing product improvement programs for assigned equipment and other functions too numerous to mention. The heavy burden imposed by these many duties simply forced the development of a data analysis technique into the background.

The Reliability and Maintainability Management Improvement Techniques (RAMMIT) Program was initiated the latter part of 1968 when the Systems Engineering Directorate at AVSCOM was directed by AVSCOM Commanding General, Major General John Norton, to evaluate an unsolicited proposal to modify aircraft systems currently in the Army inventory. The RAMMIT system was designed to process TAERS maintenance action data and other data records available to AVSCOM for the purpose of presenting it as useful information that could be used as an aid in decision-making with regard to Army aircraft and support equipment. RAMMIT has been used in data gathering, but management has not yet utilized this data to its utmost.

The current decision process for manual processing of EIRs is shown in Figure 1. This process runs into difficulty on two points. The first is due to the large number of EIRs being sent in from the field. Their number is so great that they simply cannot all be processed with the current resources. The second point concerns the human element. A great deal of "judgment and experience" is used in the decision process and the nature of this will vary from person to person. In addition, pressure is sometimes applied by outside users to influence the disposition of a particular EIR. These factors can result in inconsistent treatment of similar EIRs. The decision model simply amounts to quantifying and computerizing the above process.

PROPOSED COMPUTERIZED DECISION MODEL

The computerized model is fed information from EIRs concerning the manufacturer's part number, quantity defective and time since new. All other information is accessible at AVSCOM. This model calculates and uses the four parameters - reliability, availability, inventory and cost to live with the problem - associated with the specific item of equipment, weighs and determines priority, arranges and prints the most important EIR cases first, according to their weight and in descending order. In this way, the project engineer is notified of what job is most important. As new data is put in the computer system, it updates the previous data, subsequently, giving the most important EIR case based on the latest criteria. This particular program reads in all EIRs each time an update is required. This last procedure can be modified when put to actual use in an Army installation, to store previously read and calculated data, either on tape or on disk. Furthermore, the program is structured to accommodate additional parameters objectively and quantitatively, merely by adding more subroutines.

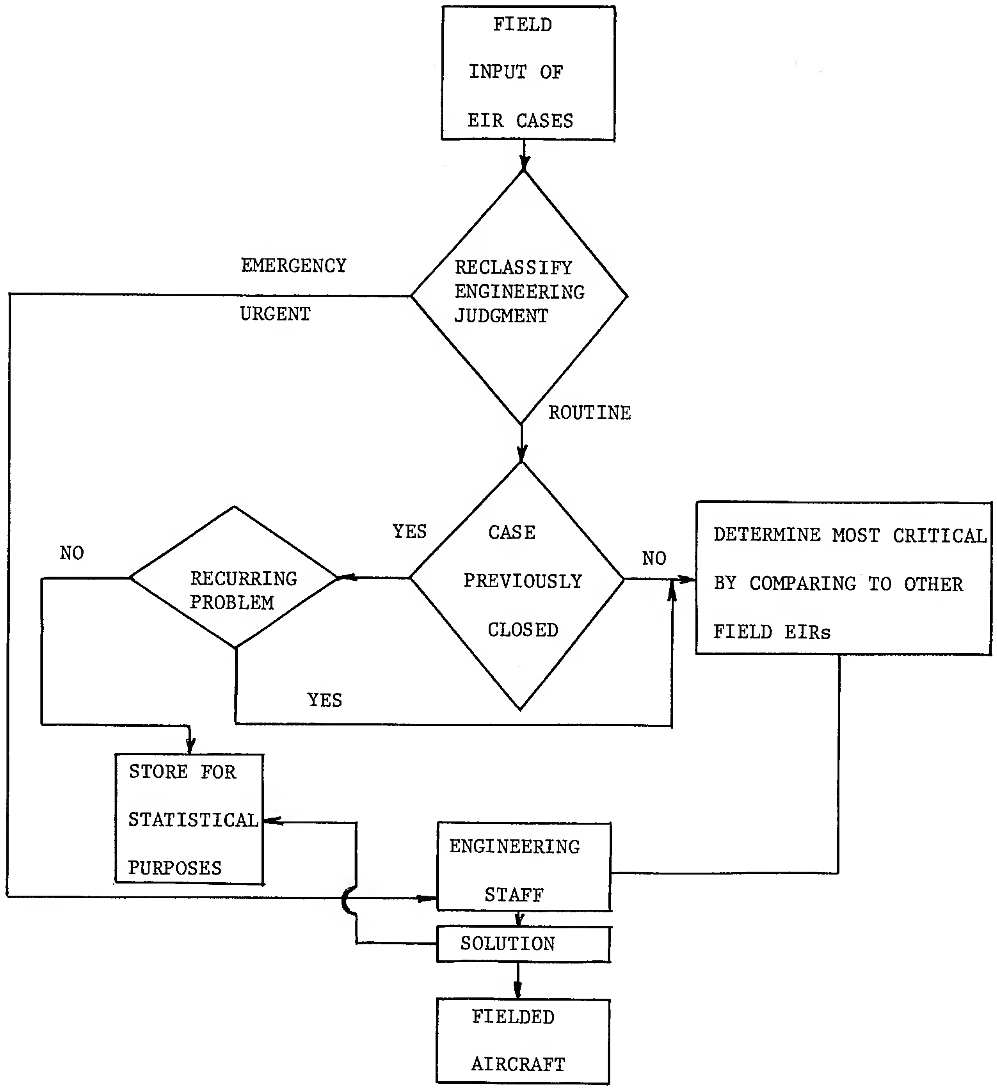


Figure 1. EIR Flow Diagram

The requirement for this proposed model or management tool is demanded by the increasing complexity and quantities of aircraft the Army has acquired. The increased feedback data, on equipment operation, maintenance and transportation, has become overwhelming in recent years; however, this tool is necessary in order to plan and manage the support system. A case study will be used to describe the computerized model. The subsystem selected as a case about which to develop the computerized decision model is part of the Army's first attack helicopter (AH-1G) system. The tail rotor subsystem was determined to be the lowest system break-down comprised of the major components shown below:

1. Quandrant Assembly
   Drawing No. 209-001-723-1

2. Cable Assembly Quadrant
   Drawing No. 209-001-728-1

3. Pulley
   Drawing No. MS202202

4. Pulley Bracket
   Drawing No. 209-001-724-1

5. Cable Assembly
   Drawing No. 205-001-724-1

6. Bracket Pulley Assembly
   Drawing No. 204-001-825-3

7. Silent Chain Assembly
   Drawing No. 204-001-739-3

Furthermore, the system was determined to be in series which is a condition where a group of components are arranged such that all must function properly for the system to succeed.

The EIR selection technique was based on assigning weight values to the four parameter values falling in certain ranges. These ranges are based on all available information. For example, reliability of a given component may be found to be above .90, and, thus, this parameter would receive a weight of 0. The four ranges and the weight values shown in Table I were arbitrarily selected by the writer based on experience.

TABLE I PRE-SELECTED PARAMETER WEIGHT VALUES

PARAMETER	RANGE	ASSIGNED NUMERICAL WEIGHT VALUE
Reliability	0-.25	1.0
	.25-.50	.75
	.50-.90	.25
	.90-1.0	0

Availability ranges and weight assignments were similarly selected. Ranges for total annual cost of inventory and total annual cost to live with the problem were selected according to AVSCOM's procurement review board dollar breakdown with weight values for each range based on the writer's experience.

The input data necessary to determine the values of these four parameters are as follows:

1. Component Number
2. Mean Time Between Failures
3. Mean Time Between Maintenance
4. Down Time
5. Rate of Demand
6. Yearly Flight Hours of Aircraft Fleet
7. Total Units Failed
8. Cost Per Unit
9. Order Quantity

The output of this program is a listing of component numbers ranked in order from the one in most urgent need of attention to the least urgent. The values of the four parameters are also given, along with additional information concerning the component. See Table II.

TABLE II COMPUTER OUTPUT

SEQ	MTBM	ACTIVE DOWN TIME	ROD	YFHOAC	TOTAL UNITS FAILED	$CPU	ORDER QUANTITY	WEIGHT
1	210	2.0	143.1	356724	239	9.33	1500	1.15
2	210	2.5	121.2	356724	16	6.19	1000	0.90
3	210	2.0	224.3	356724	4	14.76	5000	0.65
4	210	2.0	3.1	356724	5	23.43	1000	0.50
5	210	2.0	95.4	356724	3	18.03	8000	0.50

TABLE II COMPUTER OUTPUT (cont'd)

SEQ	PROJ FILE NO.	COMP NO.	RELIA	AVAIL	INVENTORY $COST	$COST TO LIVE W/PROBLEM	MTBF
1	2090017201	5	.9800	.9867	1177145.00	22424.77	148.4
2	2090017281	2	.9930	.9942	594878.31	5180.35	426.2
3	2040017393	7	.9858	.9906	445055.69	25072.59	210.0
4	202202	3	.9704	.9804	6823.10	83580.38	100.0
5	2040018253	6	.9974	.9983	154968.56	5636.92	1141.0

CONCLUSIONS

The computer model performs its evaluation task by receiving, processing, and evaluating all the failure data reported for each component or EIR case. Additionally, given values are put in from AVSCOM concerning mean down time, cost per unit, rate of demand per month, unit holding cost per month, optimum order quantity, reorder point, leadtime, cost per order, stockout costs, expected demand during leadtime greater than the reorder point, and the yearly flight hours of the aircraft fleet. The reliability for each component is calculated, weighed according to its pre-selected weight assignment values assigned a weight value and stored. The same is done for the other parameters. Thus, the management technique for determining the priority of the Army's EIR evaluation has been duplicated by the computer. The advantages of the model are numerous and of tantamount significance. One advantage is its aid in consistently processing and considering a large number of EIRs quantitatively. Total evaluation and visibility are obtained by being able to evaluate all EIRs relative to significant parameters versus conjecture and circumstantial pressures.

BIBLIOGRAPHY

1. United States Army Aviation Systems Command, Organizational, Mission and Functions of USAAVSCOM, Regulation 10-1, St. Louis, Missouri, 2 December 1968.

2. United States Army Aviation Systems Command, Organizational, Mission and Functions of USAAVSCOM, Regulation 10-2, St. Louis, Missouri, 2 December 1968.

3. Chrysler, E.J. Col., Personal Interview. August 20, 1969.

4. Department of the Army Technical Manual 55-1520-221-35P-1, Helicopter, Attack - AH-1G (Bell) Helicopter, Flight Trainer - TH-1G (Bell), Headquarters, Department of the Army, May, 1970, pp.2, 476, 264.

5. Department of The Army Technical Manual 55-1520-221-35P-2, Helicopter, Attack - AH-1G (Bell) Helicopter, Flight Trainer - TH-1G (Bell), Headquarters, Department of the Army, May 1970, pp.998, 1000.

6. Harring, M. G., and Greenman, L. R., Maintainability Engineering, prepared under contract number DA-31-124-ARO-D-100-34 with Martin-Marietta Corporation and Duke University, 1965.

7. International Business Machines, System/360 Scientific Subroutine Package (360-CM-03X) Version III, Programmers Manual.

8. University of Missouri-Rolla Computer Center, WATFIV/360 Notes Rolla, Missouri, 1971.

9. Lloyd, D. K., and Lipow, M., Reliability: Management, Methods, and Mathematics, Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1962.

10. Hiller, Fredrick S., and Lieberman, Gerald J., Introduction to Operations Research, San Francisco, Cambridge, London, Amsterdam: Holden-Day, Inc., 1969.

11. McMillan, Calude and Gonzales, Richard F.,Systems Analysis, Nobleton, Ontario: Irwin-Dorsey Limited, 1968.

DC9-30 REFRIGERATION SYSTEM DIAGNOSIS BY COMPUTER

J. Albert, Project Engineer,
Eastern Air Lines, Miami, Florida

Summary

A novel method of quickly diagnosing DC9-30 refrigeration problems is presented. Readily obtainable steady state data are substituted in mathematical models relating to component performance, and the results compared to prescribed operating limits by computer for instantaneous diagnosis. In addition, hot day conditions are mathematically simulated and potential system problems predicted to provide preventative maintenance. An "on condition" maintenance programme based on these concepts has been implemented, resulting in significant reduction in maintenance costs and increased system reliability.

This instantaneous diagnostic approach to aircraft maintenance practices is believed to be the first of its type in the airline industry.

Introduction

During normal aircraft operation, air conditioning components degrade until adequate cabin cooling is not provided. Maintenance programmes based on "hard time" for the individual components have been ineffectual in maintaining an acceptable level of system reliability because of the tenuity of most of the component failure time relationships. Furthermore, existing trouble-shooting techniques have been inadequate in quickly isolating problems, resulting in prolongation of problems and high maintenance costs.

Because of the concern to ensure adequate passenger comfort and to reduce system maintenance costs, it was concluded that a preventative maintenance programme was required to predict system deficiencies, featuring a simple "in situ" test, compatible with the daily airline operation time framework, together with instantaneous system diagnosis.

System Description

The Douglas DC9-30 refrigeration system is designed and built by AiResearch and has proved to be a reliable and well designed system.

The system consists of two identical air cycle systems supplied with pneumatic air by either the engines or the auxiliary power unit (APU). A simplified system schematic appears in Figure 1. Heat rejection is effected by an electric fan on the ground and ram air in flight. Overheat protection is provided by thermal switches at the pack outlet and compressor discharge while turbine overspeed is prevented by a thermal switch at the turbine inlet. Cockpit indication is provided for pneumatic supply pressure, regulated supply pressure and pack discharge temperature.

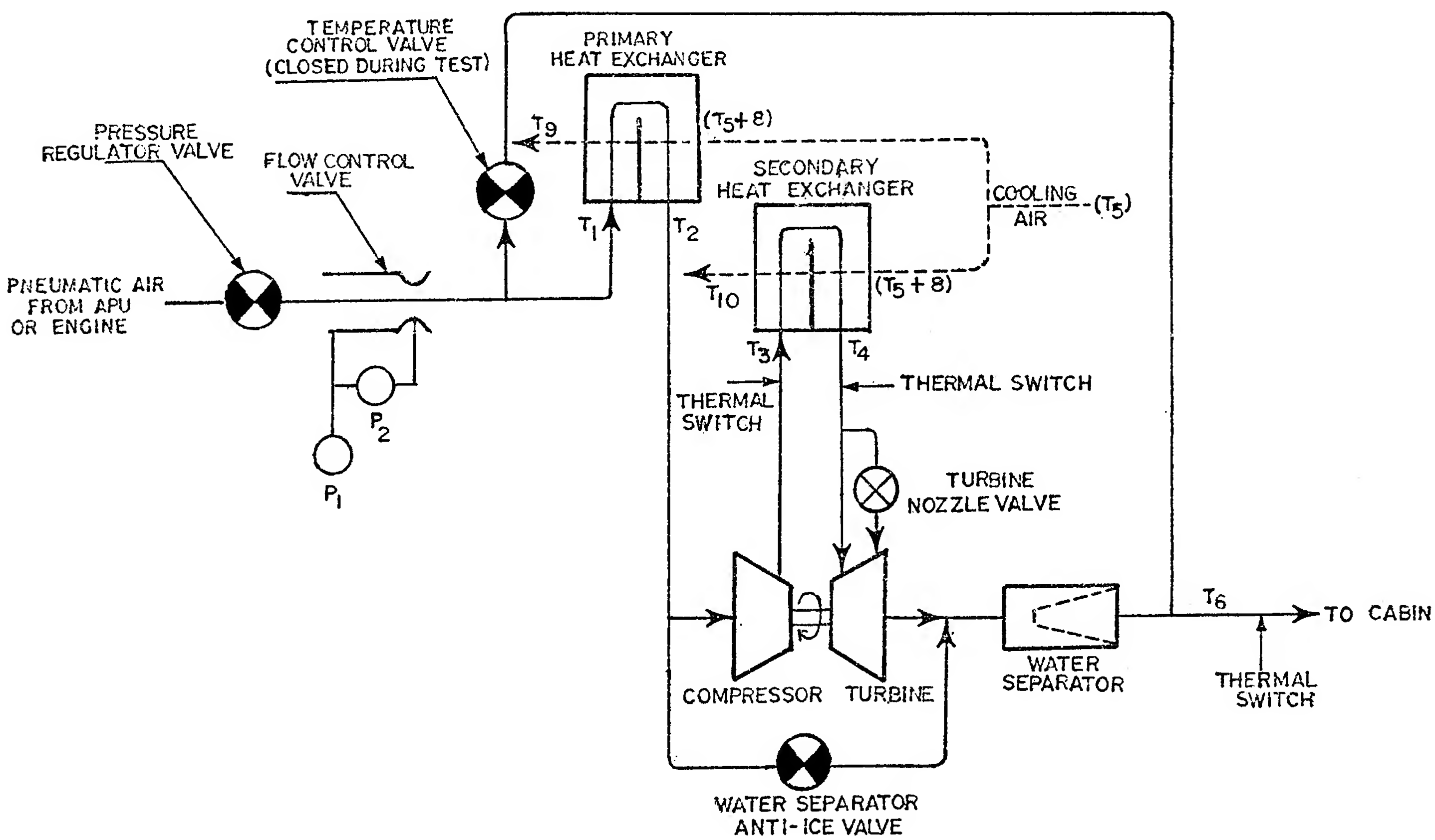


FIGURE I: SCHEMATIC OF SIMPLIFIED OF DC9-30 AIR CYCLE SYSTEM

System Malfunctions

The major causes of system malfunction are:

(a) Low air flow caused by malfunctioning pressure regulator and/or flow control valve. Although the heat exchangers have less air to cool, the turbine work and heat of compression are reduced, resulting in a high turbine discharge temperature and poor cooling. Figure 2 provides a graphical representation.

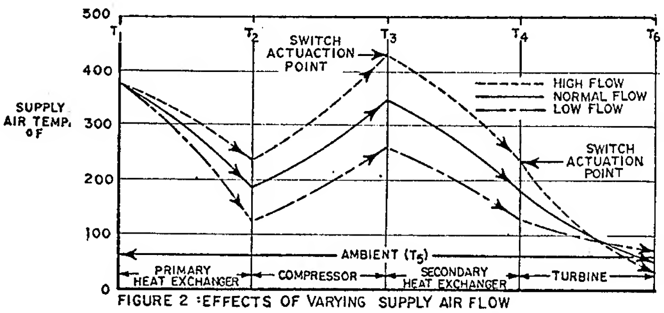


FIGURE 2 :EFFECTS OF VARYING SUPPLY AIR FLOW

(b) High supply air flow also caused by malfunctioning pressure regulator and/or flow control valve. Excessive supply air produces increased turbine work and heat of compression with actuation of either the compressor discharge or turbine inlet thermal switch, resulting in pack shutdown. A graphical representation is also shown in Figure 2.

(c) Degraded air cycle machine as a result of turbine nozzle and blade erosion. A reduction in turbine work and, in turn, heat of compression results in high turbine discharge temperature and poor cooling. Figure 3 shows the effects of a degraded air cycle machine.

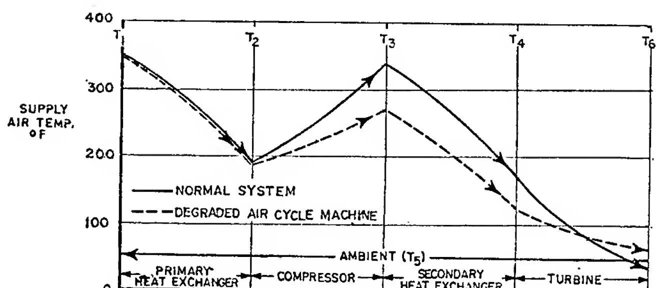


FIGURE 3 :EFFECTS OF DEGRADED AIR CYCLE MACHINE

(d) Premature opening of the water separator anti-ice valve. Relatively hot supply air from the primary heat exchanger outlet is allowed to mix with the relatively cool turbine discharge air, resulting in poor cooling. A graphical representation is shown in Figure 4.

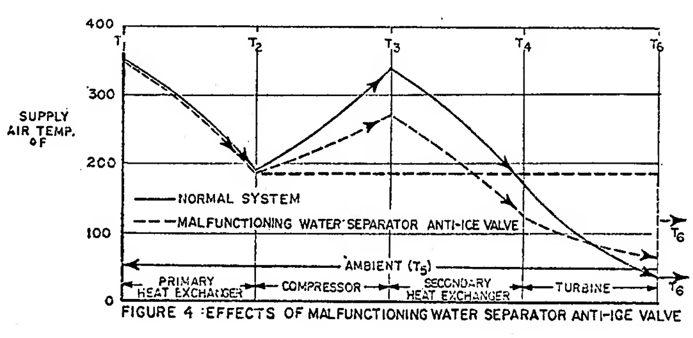


FIGURE 4 :EFFECTS OF MALFUNCTIONING WATER SEPARATOR ANTI-ICE VALVE

(e) Degraded primary and/or secondary heat exchangers. Degraded heat exchangers reduce heat rejection and results in a high turbine discharge temperature and poor cooling. In extreme cases of degradation, actuation of either the compressor discharge or turbine inlet thermal switch will occur, resulting in pack shutdown. Figure 5 represents the effects of degradation in both units without actuation of either thermal switch.

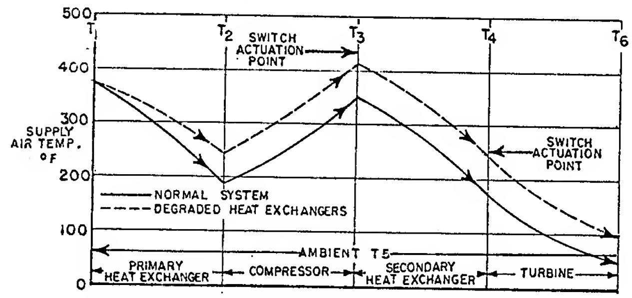


FIGURE 5: EFFECTS OF DEGRADED HEAT EXCHANGERS

Programme Development

The malfunctions described produce 41 combinations of poor system operation. Although a single component can be changed at a planned time, restoration of system performance does not necessarily follow, because of the thermodynamic balance between the components.

The object of any system maintenance is that, in addition to the integral parts, the sum of the integral parts (i.e. the system) shall perform within prescribed limits of operation and reliability.

Cockpit indications provide diagnosis of only the pressure regulator valve by means of the regulated supply air pressure indicator. The pneumatic supply pressure indication is not a function of system performance while the pack discharge temperature, though indicating total system performance, provides no fault isolation.

A review of the system installation showed that all parameters for complete system diagnosis could be readily measured, involving modest labour and material resources.

AiResearch data were obtained, from which mathematical models were developed for the performance of each component, independent of ambient conditions.[1] With the efficiency of each component determined, the aggregate system performance can be predicted for any ambient. Hence, the refrigeration system can be tested in the spring and system performance predicted under summer conditions. Defective components can be detected and replaced before summer, providing a preventative maintenance programme. An annual fleetwide pre-summer efficiency was to form the basis of such a programme.

A performance standard based on Eastern's most thermally severe station (Dallas) was chosen, requiring the cabin to be cooled with the APU source to a minimum comfort level from hot soaked conditions in an acceptable time.[2] In addition, the overheat switches were not to actuate with the engine at take-off power at the same selected ambient. Fuselage thermodynamic data from Douglas enabled the cabin cooling requirements for the selected hot day conditions to be determined.[3]

Because of the complexity of the calculations and logic process, it was decided to computerize the entire procedure to provide instantaneous diagnosis. It was determined that the existing communications terminals could be used for transmitting test data and receiving the required instantaneous information.

Test Procedure

Heat exchanger supply air inlet and outlet temperatures are obtained by removing plugs and switches from the heat exchanger inlet and outlet ducts and installing dial temperature gauges with stainless steel adapters. Heat exchanger cooling air outlet temperatures cannot be readily measured. Flow control valve pressures are measured with pressure gauges connected to the valve by flexible hoses replacing the rigid valve sense lines as shown in Figure 6.

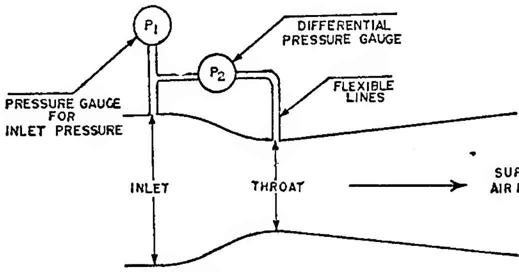


FIGURE 6: FLOW CONTROL VALVE SCHEMATIC

With APU pneumatic source, each pack is allowed to thermally stabilize with the temperature control valve closed to ensure that the entire air supply passes through the pack, and the turbine nozzle valve open to allow the flow control valve to operate at its regulation point. After stabilizing, pressure and temperature gauge readings are recorded. Ambient conditions are obtained during pack stabilization. The data points are arranged in specific order and relayed to the computer through a communications terminal. The data are automatically processed and system diagnosis together with defective components instantaneously displayed.

Testing of both packs requires 2 hours (4 manhours). One set of equipment for testing both packs simultaneously costs approximately $450.

Computer Details

IBM 2740 or UNIVAC U100 terminals, located at all stations, are used for transmitting data to an IBM 360/65 computer in Miami. The "real time" programme is contained in three 10K modules written in A/L. Modified IBM Fortran macro routines are used for computing exponential functions.

Construction of Mathematical Models

The following simplified constructions cover the major features of the programme. Duct temperature losses are neglected for further simplification of the presentation.

Test data are designated as follows (Figure 1 also refers). All pressures are in psig and all temperatures in °F.

- P_1 = flow control valve inlet pressure
- P_2 = flow control valve inlet-to-throat differential pressure
- T_1 = primary heat exchanger supply air inlet temp.
- T_2 = primary heat exchanger supply air outlet temp
- T_3 = secondary heat exch. supply air inlet temp.
- T_4 = secondary heat exch. supply air outlet temp.
- T_5 = ambient temp.
- T_6 = pack discharge temp.
- R = relative humidity (%)

Flow Control Valve (Venturi Type)

Applying the formula for compressible flow through a venturi with valve inlet and throat dimensions known[4], supply air in lb/min. is given by:

$$W = 336.45(P_1+14.7)\left(\frac{P_1+14.7-P_2}{P_1+14.7}\right)^{0.714} \times \left\{\frac{\frac{1}{T_1+464}\left[1-\left(\frac{P_1+14.7-P_2}{P_1+14.7}\right)^{0.286}\right]}{1-0.16944\left(\frac{P_1+14.7-P_2}{P_1+14.7}\right)^{1.43}}\right\}^{0.5}$$

Pressure Regulator Valve

Flow control valve inlet pressure (P_1) is a direct indication of regulated supply pressure.

Water Separator Anti-Ice Valve

Let T_7 = calculated turbine discharge temp.
T_8 = turbine discharge temp. (dry air rated)
H = air moisture content

To determine the bypass air flow through the water separator anti-ice valve, calculated turbine discharge temperature (T_7) and primary heat exchanger supply air outlet temperature (T_1) are weight averaged and equated with the pack discharge temperature (T_6).

To determine turbine discharge temperature (T_7), dry air rated turbine discharge temperature (T_8) is determined by subtracting the dry air rated turbine temperature drop from the turbine inlet temperature (T_4). The water content of the air (H) is computed from relative humidity (R) and ambient temperature (T_5). Referring to the graphical representation of the process in Figure 7, the values of H and T_8 determine the enthalpy of the turbine discharge air. Maintaining constant enthalpy at the design turbine discharge pressure at saturation conditions, turbine discharge temperature can be determined.

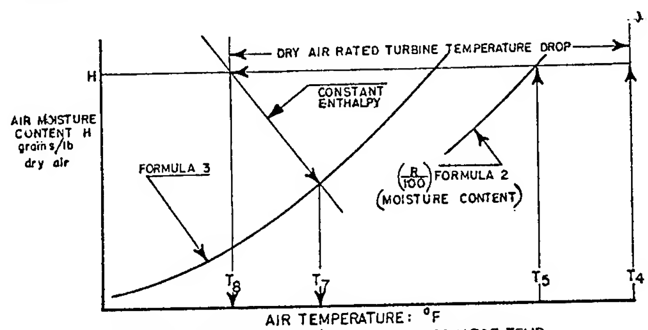


FIGURE 7: DETERMINATION OF TURBINE DISCHARGE TEMP.

The mathematical model for computing turbine discharge temperature is developed as follows, with equations first developed for saturated air:

$$H = \left(\frac{T_5+460}{416.1}\right)^{19.41} \quad \text{at } 29.92\text{" HgA} \qquad (2)$$

$$H = \left(\frac{T_7+460}{421.2}\right)^{19.44} \quad \text{at } 38\text{" HgA} \qquad (3)$$

where 38" HgA is design turbine discharge pressure with turbine nozzle valve open.

Also, the coefficient of specific heat for air at constant pressure (C_p) was derived as follows:

$$C_p = 0.24 + 0.00003\,\frac{R}{100} \quad \text{(Formula 2)} \qquad (4)$$

For air at constant enthalpy:

$$T_7 \text{ or } T_8 = 0.6417H + \text{constant}$$

$$T_7 + 0.6417(\text{Formula 3}) = T_4 - 1.039(T_3 - T_2) + 0.006417R(\text{Formula 2}) \qquad (5)$$

where 1.039 is reciprocal of air cycle machine efficiency of 0.96

Formula 5 is solved for T_7 (turbine discharge temperature).

Therefore, supply air flow through the air cycle machine and secondary heat exchanger when bypassing equals:

$$(\text{Formula 1})\left[1 - \left(\frac{T_6 - T_7}{T_2 - T_7}\right)\right] \qquad (6)$$

Primary Heat Exchanger

Assuming zero convection and radiation:

$$\text{Heat Transfer}(Q) = W.C_p(T_1 - T_2)$$

$$= (\text{Formula 1})(\text{Formula 4})(T_1 - T_2)$$

$$= W_f(\text{Formula 4})(T_9 - T_5 - 8) \qquad (7)$$

where W_f = equivalent cooling air flow
T_9 = equivalent cooling air outlet temp.
8 = temperature rise of cooling air due to fan

Logarithmic Mean Temperature Difference (LMTD) =

$$\frac{(T_1 - T_9) - (T_2 - T_5 - 8)}{\text{Log}_e\dfrac{(T_1 - T_9)}{(T_2 - T_5 - 8)}} \qquad (8)$$

A graphical representation of the heat transfer process is shown in Figure 8.

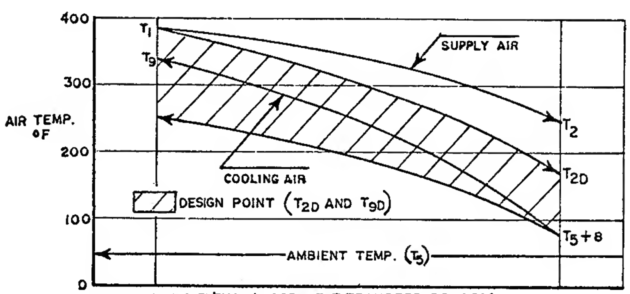


FIGURE 8: HEAT EXCHANGER HEAT TRANSFER PROCESS

From AiResearch steady state data[1], the following relationship was derived:

$$UA = W^{0.3867} \times W_f^{0.3268}$$

$$= (\text{Formula 1})^{0.3867} \times W_f^{0.3268} \qquad (9)$$

When W_f is at the design point (79.2 lb/min)

$$UA = 4.17(\text{Formula 1})^{0.3867} \ldots\ldots (UA_1)$$

$$\text{Also LMTD} = \frac{Q}{UA}$$

Therefore, when W_f is at the design point:

$$(\text{Formula 8}) = (\text{Formula 4})(T_1 - T_2) \times (\text{Formula 1})^{0.6133} \qquad (10)$$

Formula 10 is solved for T_9 (cooling air outlet temperature), which when substituted in formula 7, the equivalent cooling air flow (W_f) is computed to produce an actual heat transfer of Q. With equivalent cooling air flow normalized to hot day conditions, and an assumed hot day supply air flow substituted in formula 9, the primary heat exchanger UA factor under hot day conditions (UA_2) is computed. Figure 9 is a graphical representation of the process.

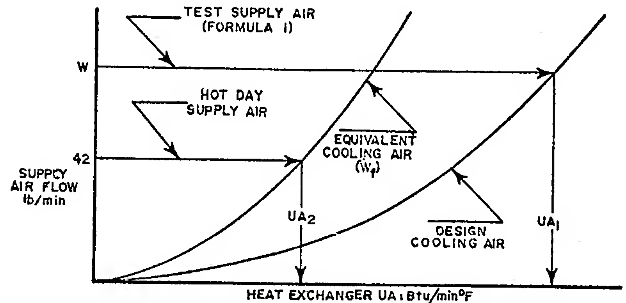


FIGURE 9: DETERMINATION OF HEAT EXCHANGER UA FACTOR

Secondary Heat Exchanger

Assuming no bypass air and zero convection and radiation:

$$\text{Heat transfer}(Q) = W.C_p(T_3 - T_4)$$

$$= (\text{Formula 1})(\text{Formula 4})(T_3 - T_4)$$

$$= W_f(\text{Formula 4})(T_{10} - T_5 - 8) \qquad (11)$$

where W_f = equivalent cooling air flow
T_{10} = equivalent cooling air outlet temp.
8 = temperature rise of cooling air due to fan

Logarithmic Mean Temperature Difference (LMTD) =

$$\frac{(T_3 - T_{10}) - (T_4 - T_5 - 8)}{\text{Log}_e\dfrac{(T_3 - T_{10})}{(T_4 - T_5 - 8)}}$$

Figure 7 provides a similar qualitative representation of the heat transfer process in the secondary heat exchanger.

From AiResearch steady state data[1], the following relationship was derived:

$$UA = W^{0.5599} \times W_f^{0.2222}$$

$$= (\text{formula 1})^{0.5599} \times W_f^{0.2222}$$

When W_f is at the design point (106.3 lb/min)

$$UA = 2.82(\text{Formula 1})^{0.5599} \ldots\ldots\ldots (UA_1) \qquad (13)$$

$$\text{Also LMTD} = \frac{Q}{UA}$$

Therefore, when W_f is at the design point:

$$(\text{Formula 12}) = (\text{Formula 4})(T_3 - T_4) \times (\text{Formula 1})^{0.4401} \qquad (14)$$

Formula 14 is solved to T_{10} (cooling air outlet temperature), which when substituted in formula 11, the equivalent cooling air flow (W_f) is computed to produce an actual heat transfer of Q. With equivalent cooling air flow normalized to hot day conditions, and an assumed hot day supply air flow substituted in formula 13, the primary heat exchanger UA factor under

hot conditions (UA_2) is computed. Figure 9 also represents the secondary heat exchanger heat transfer process.

If air is bypassed through the water separator anti-ice valve, as determined in formulae 2 - 6, formula 6 is substituted for formula 1.

Air Cycle Machine

From AiResearch steady state data[1], the following relationship was derived for the air cycle machine with the turbine nozzle valve open and no bypass air:

$$W_c = (0.9747W)^{1.8265}$$
$$= \left[0.9747(\text{Formula } 1)\right]^{1.8265} \qquad (15)$$

where W_c = design compressor work rate (Btu/min) at flow W

Actual work done at flow $W = W.C_p(T_3 - T_2)$

$$=(\text{Formula } 1)(\text{Formula } 4) \times (T_3 - T_2) \qquad (16)$$

Therefore efficiency(E) $= \dfrac{\text{Formula 16}}{\text{Formula 15}}$

$$= 1.048(T_3 - T_2)(\text{Formula } 4) \times (\text{Formula } 1)^{-0.8265} \qquad (17)$$

If air is bypassed through the water separator anti-ice valve, as determined in formulae 2 - 6, formula 6 is substituted for formula 1.

Figure 10 provides a graphical representation of the air cycle machine efficiency computation.

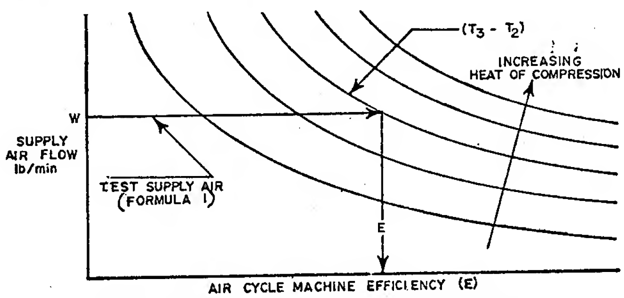


FIGURE 10: DETERMINATION OF AIR CYCLE MACHINE EFFICIENCY

By correlating steady state data with the air cycle machine overhaul test requirements[5], the following relationship was derived for the air cycle machine with the turbine nozzle valve closed and no bypass air:

$$W_t = (1.255W)^{1.8265}$$
$$= \left[1.255(\text{Formula A})\right]^{1.8265}$$

where W_t = design turbine work rate (Btu/min) at flow W

System Performance Standard

The following relationship was developed from Douglas data[3] for the DC9-30 aircraft under the following initial hot soaked conditions:

Ambient: 100°F dry bulb, 80°F wet bulb
All doors closed
No electrical load
No passengers
Daylight - sunny

Initial Hot Soaked Conditions (Continued)

APU operation
Water separator anti-ice valve closed
Turbine nozzle valve closed
Supply air 84 lb/min (42 lb/min per pack)
$T_1 = 416°F$

$$e^{-2.117T} = \frac{1209T_{13} + 36}{1209T_{13}} \qquad (19)$$

where T_{13} = turbine discharge temperature
T = pack operating time (hours)

The minimum comfort level for transient occupancy at 100°F ambient (the Dallas design point) is 83°dry bulb and 50% relative humidity[2].

A relationship between pack operating time (T) and turbine discharge temperature (T_{13}), to produce a cabin temperature of 83°F under the selected hot day conditions derived from formula 19, is shown in Figure 11.

With the following values established for the primary heat exchanger under hot day conditions:

UA factor from formula 9
Equivalent cooling air flow from Formula 7 and 10
Supply air flow 42 lb/min (assumed)
Supply air inlet temp. = 416°F
Cooling air inlet temp. = 108°F

primary heat exchanger supply air outlet temperature (T_{11}) is computed from formula 8 and 10 suitably modified.

For a 100% efficient air cycle machine with the turbine nozzle valve closed, the heat of compression at 42 lb/min supply air flow (from formula 18) equals

$$\frac{0.96(\text{Formula 18})}{42.C_p} = 131.4°F$$

where 0.96 = mechanical efficiency of air cycle machine
C_p = specific heat of air at hot day conditions (0.2437 Btu/lb°F)

Therefore, for any air cycle machine efficiency (E), secondary heat exchanger supply air inlet temperature equals:

$$T_{11} + 131.4(\text{Formula 17})$$

With the following values established for the secondary heat exchanger under hot day conditions:

UA factor from formula 13
Equivalent cooling air flow from formulae 11 and 14
Supply air flow 42 lb/min (assumed)
Supply air inlet temp. = T_{11} + 131.4 (Formula 17)
Cooling air inlet temp. = 108°F

secondary heat exchanger supply air outlet temperature (T_{12}) is computed from formula 12 and 14, suitably modified.

By substitution in formula 18, the dry air rated turbine temperature drop at 42 lb/min supply air flow, with the turbine nozzle valve closed, is 138.7°F. Under these conditions, the following model was developed:

$$H = \left(\frac{T_{13} + 460}{416.1}\right)^{19.05} \qquad (21)$$

Following a similar procedure for determining test turbine discharge temperature (T_7):

$$T_{13} + 0.6417(\text{Formula 21}) = (\text{Formula 20}) - 138.7(\text{Formula 17}) + 78.3 \qquad (22)$$

where T_{13} = predicted turbine discharge temperature under hot day conditions.

From the above hot day formulae 19 - 22, the relationship of predicted turbine discharge temperature and air cycle machine efficiency, with constant primary and secondary heat exchanger efficiency of 95% of design, can be derived as shown in Figure 11. The 95% efficiency factor allows for the inability of overhauled heat exchangers to be totally restored to the design condition.

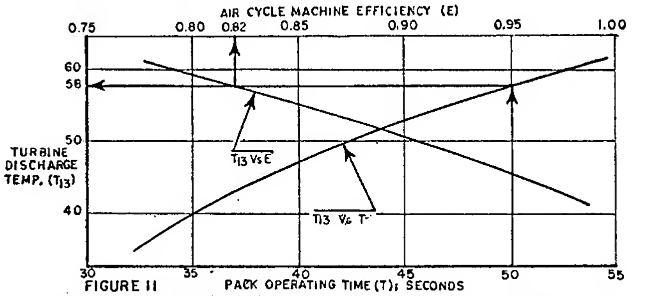


FIGURE 11

With an operating time standard of 50 minutes, an air cycle machine efficiency of 0.82, with a turbine discharge temperature of 58°F, is required to provide the minimum comfort level of 83°F. (Fig. 11 refers).[6] Referring to AiResearch data for the water separator[6], sufficient water is extracted under the hot day conditions to produce a relative humidity of less than 50%, thereby complying with the minimum transient comfort levels[2].

For the hot day take-off power conditions, applicable flow and temperatures were substituted in the hot day formulae. As minimum compliance with the cabin cooling standard does not incur actuation of either the compressor overheat or turbine overspeed switch, the programme is confined to simulating the hot day APU operation. A graphical representation of the hot day pack performance is provided in Figure 12.

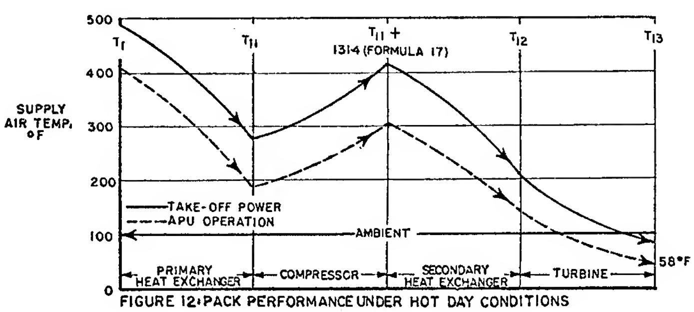


FIGURE 12: PACK PERFORMANCE UNDER HOT DAY CONDITIONS

Flow Chart

The data are first tested for validity by applying prescribed limits to each data point. Any data point exceeding its limit is automatically displayed. In order to detect data transpositions, the data must show that:

a) Heat exchanger supply air inlet temperature exceeds the outlet temperature.

b) Heat exchanger supply air outlet temperature exceeds the cooling air inlet temperature.

When transposing occurs, the offending data points are automatically displayed.

If the test data are in order, the following simplified computation procedure is automatically followed. Components out of operating limits are automatically displayed on the print-out.

Flow Control Valve

The computed flow in formula 1, normalized to hot day conditions, is compared to the valve operating limits of supply air flow.

Pressure Regulator Valve

The regulated pressure (P_1) is compared to the valve operating limits of supply air pressure.

Water Separator Anti-Ice Valve

If the valve is bypassing supply air, as determined in formula 2 - 6, and pack discharge temperature (T_6) exceeds 40°F (allowing for permissable valve leakage), the water separator anti-ice valve is defective.

Air Cycle Machine

If the air cycle machine efficiency (E), computed in formula 17, is below 0.82 (Figure 11 refers), the unit is below acceptable performance.

Heat Exchangers

If the air cycle machine is below acceptable performance (less than 0.82 efficiency), an air cycle machine efficiency of 1.0 is substituted in formulae 20 and 21 to simulate an air cycle machine change, whereas the computed efficiency (E) is applied if the unit is in order. If the predicted turbine discharge temperature (T_{13}) exceeds 58°F (Figure 11 refers), both primary and secondary heat exchangers require changing. As both units are in a common plenum, the effect of changing one unit cannot be predicted, as blockage in one unit produces an exaggerated efficiency in the other.

Results of Programme Implementation

Implementing the programme in April 1972 resulted in a reduction of 60% in the DC9-30 refrigeration log report rate and a reduction of 31% in the applicable component removal rate, compared to the equivalent 1971 period. A greater improvement is anticipated in future years, as all corrective actions arising from the tests were not completed before summer, as planned, because of a late start.

References

1: DC9-39 Air Conditioning Half System Steady State Data: AiResearch Report 67-2059.

2: Marks' Mechanical Engineers' Handbook - Air Conditioning Summer Design Conditions - pages 12-120.

3: Douglas Letter CI-781-M-1818/LSS: "DC-8 and DC-9 Ground Air Conditioning Requirements" dated 6/2/69.

4: AiResearch Drawing 396050: DC9-30 Flow Control Valve.

5: AiResearch Overhaul Manual 204950: DC9-30 Cooling Turbine.

6: AiResearch DC-9 Water Separator Performance Chart 179460-7-1.

Acknowledgements

To Doug Bell and Danny Koren of Eastern Air Lines' Computer Sciences for translating the mathematical models and flow chart into a workable computer package.

A COMPUTER SIMULATION OF THE LIFE CYCLE

Virgil Rehg
U.S. Air Force Institute of Technology
School of Systems and Logistics
Wright-Patterson AFB, Ohio

Summary

Presented in this paper is a consideration of the affects of early decisions in equipment acquisition upon the life cycle costs of the equipment. It describes and illustrates how a stochastic simulation model for an equipment or system can be designed. It also illustrates the outcomes in terms of cost and delays that result when the system is simulated overtime for a given set of failure rates, repair times and other decision parameters.

The factors that are included in computing the life cycle costs are discussed and the procedure for building a simple model is also presented. Flow charts are used to illustrate the logic of the decision points that must be considered in constructing the model. Random number generators are explained and then used to produce equipment failure times, repair times, equipment check-out results after repair, lead times for inventory replacement, personnel availability and other related stochastic processes. A simple model is operated so that an example of the analysis can be demonstrated.

Introduction

At the meeting of the symposium in 1969, a paper entitled "Reliability Management Simulation Exercise" was presented. It described a reliability simulation game in use at the Air Force Institute of Technology, School of Systems and Logistics as part of a course in reliability. In this exercise the players are placed in an environment resembling the real world and must make decisions, trade-offs and solve problems that arise in the acquisition of a system. When the exercise is completed, the students are fully aware of the critical role that reliability and maintainability play in system acquisition and the costs that are expected during the life cycle of the equipment.

This paper is a follow-up to the first paper and considers in more detail the affects that decisions made early in the life cycle have upon costs experienced later in the life cycle. In this paper, a Monte Carlo Simulation (or stochastic simulation) model is presented that permits the student to "live" through the life cycle of an equipment while the equipment is in the research and development stage. The advantage of such an experiment is that various trade-offs can be tested and evaluated before a final decision is made. Since parameters that are important to both the buyer and seller can be tested, the model is useful to both parties involved in a contract.

The objective of this paper is to present a method for performing this type of an analysis and provide the reader with an example of how the MTBF (Mean Time Between Failure) affects LCC (Life Cycle Costs). The paper does not consider all the life cycle factors that are possible because some factors are peculiar to only certain applications. Emphasis is placed on the LCC-MTBF relationship and the common costs that can be experienced in operating equipment, on the use of random number generators, and the collection of data that is needed to calculate the life cycle costs.

The Monte Carlo Simulation should not be confused with the simulation game. The Monte Carlo Simulation is part of the simulation game for use at the players discretion. In the Monte Carlo Simulation, a system is operated or tested over a period of time so that the user can judge the expected outcome of his decisions, and if necessary, repeat the process using the same or other parameters. In the simulation game, the players are provided with information and various results in a real time mode. In this paper our attention will be concentrated on the Monte Carlo Simulation.

To illustrate the logic in building such a model, tree diagrams and flow charts are used. To produce random events including such things as failure times, personnel availability, human factors problems, repair times, and other events of a stochastic nature, random number generators are used.

The paper will be considered successful if the reader, not already familiar with Monte Carlo Simulation and/or life cycle costs, gains an insight into the importance of the MTBF-LCC relationship, and the procedures for constructing a Monte Carlo simulation model. To construct a model, we must begin by defining the problem. Are we trying to find what MTBF to use in a design? Or perhaps we wish to know what affect priority repair jobs have on the operation of a system. We must then collect information on the factors to be considered and data on parameters to be used in the simulation. Then we can build the mathematical model and test it by making a short run to determine its validity. Now we can incorporate the model into a computer program, run the program and perform an analysis on the output.

Factors that Affect Life Cycle Costs (LCCs)

Let us proceed by listing some of the factors that are called life cycle costs and isolating those that are part of this study. They are called LCCs because they are the costs that will be experienced during the life of the equipment.

Maintenance	Waiting Time
Spares	Transportation
Test Equipment	Technical Data
Training	Operating Costs
Inventory Management	Facilities
Downtime	Installation

There may be more, but this list should give the reader some idea as to the type normally considered. In this paper we shall consider <u>maintenance</u>, <u>spares</u>, <u>test equipment</u>, <u>inventory management</u>, <u>downtime</u> and <u>waiting time</u> because they are common to most equipments, they constitute the largest part of the LCC dollar and also because most of them are dependent upon the MTBF and/or MTTR, (Mean Time to Repair).

Maintenance

The costs of maintenance occur when equipment fails.

They include direct labor and overhead. This cost can be computed by multiplying the time it takes to repair the equipment by the cost of labor and summing over the number of failures. It can also be calculated by generating failure times and repair times and computing the labor cost for each failure and then adding them.

Test Equipment

This factor represents the cost of test equipment that is necessary to analyze equipment failures. For electronic equipment the test equipment could be the various meters used for testing. For mechanical equipment it could be gauges, tools, meters, or a combination of all three.

Training

To perform an analysis of an equipment failure, the individual must be trained; when modifications are made retraining may be necessary; and because of personnel turnover, training must take place periodically. There also may be different levels of training required for maintenance. In practice, training is a never ending cost and must be recognized, but we shall not consider it at this time even though it is part of the simulation game.

Inventory Management

The cost of managing an inventory depends upon the number and quantity of items in the inventory. In turn, the number and quantity depend upon the failure rate of the equipment, the life of replacement item and the length of time it takes to replace the item.

Spares

The number of spare items needed depends upon the failure rate of the items. The cost of spares is found by multiplying the number of spares required by the price of the spare. Since the number of spares required depends upon the number of failures, spares are a function of the MTBF.

Downtime

This is the length of time it takes to repair the equipment, hence, it is dependent upon the MTTR, the availability of test equipment and maintenance men and whether the repair facilities are busy or idle when the failure occurs.

Waiting Time

Waiting time is usually part of downtime but has been isolated here so that we can illustrate the relationship between waiting time and the MTBF and MTTR. It is defined as the time an equipment must wait for service, wait for a replacement part, wait for transportation or whatever the causes of waiting might be.

Transportation

The cost of transportation includes costs incurred to ship the equipment to its intended location. However, the largest part of the transportation experienced over the life cycle is the cost of shipping the equipment and/or replacement parts for equipment when it has failed. Since the number of failures depends upon the failure rate, the transportation cost is in part a function of the failure rate.

Technical Data

Technical data includes all types of specifications, standards, drawings, instructions, manuals, test results and reports used in all stages of the life cycle.

Operating Costs

The costs peculiar to the operation of an equipment are the subject of this factor. This includes such things as fuel, power, manpower, depreciation and those costs normally associated with the operation of an equipment.

Facilities

The facilities are the buildings and space required to house the equipment and its supporting elements. If new buildings must be constructed for the equipment and office space for the operators and if both must be air conditioned, then the costs associated with them would be considered facility costs.

Installation

Many times an equipment requires installation and check-out by specialists. If so, the related costs are installation costs. These brief descriptions should make the reader aware of what is meant by an LCC factor. As stated earlier, you may think of several more that should be included for certain equipments. The list presented is not intended to be complete, but is indicative of the concept being considered.

Building the Model

Let us begin this section of the paper by considering a simple example to illustrate the concept to be presented later. In the example, we shall ask ourselves a series of questions that in real life would be apparent when the problem arises. We ask them here because they represent the type of elements that must be considered in building a Monte Carlo Simulation.

Example 1

The situation in the example deals with the failure of a light bulb in our kitchen at home. The questions we usually consider are:

1. How do we know the light has failed? Obviously, if it is a single light then we are in darkness and cannot see what we want to see. But if it is a redundant system with two light bulbs or multiple light bulbs, we may not even be aware of the fact that one bulb has failed. It may be necessary to design the equipment with some sort of test equipment built into the system that registers every failure. In the construction of the Monte Carlo model, we must consider these factors because of the affect on the failure rate, repair time and number of spares required.

2. Is a special tool required to remove and replace the bulb? Is it available and does it work? By this I mean, is a step ladder needed? Screwdriver? Pliers? Special bulb removing tool? Or can it be done by hand? In the Monte Carlo Simulation, we must consider the probability of special tools being available and the probability that they are in operating condition. If they are not available or do not work then waiting time will be experienced until a substitute is found or the tools become available in

working condition.

3. Does the replacement operation require a special skill? To change a light bulb, probably not. But suppose the light bulb could not be reached until some other things were removed. For example, suppose the light bulbs were in an electric dryer or a car radio. Not everyone could make such a replacement. In the Monte Carlo Simulation, we must consider the availability of personnel in such a simulation.

4. Must other units be shut down to make the repair? Some people might turn off all the power in the house to make such a repair. This will not cause a problem other than with an electric clock which now must be reset. But consider the complex equipment in many systems that cannot be shut down without causing lengthy delays in restarting and perhaps causing other failures.

5. Do we have a replacement bulb in the house? (This may have been one of the first questions we literally asked ourselves). If we do not, then we again experienced both downtime and waiting time. In fact, it may be very costly because we may have to go out to eat when the kitchen light is burned out. Whether we have a light depends upon how well we are managing the light bulb inventory. In the Monte Carlo Simulation, it is necessary to check the inventory status, to deduct an item from inventory when a replacement is made and to replenish the inventory at the appropriate time.

6. Questions two and three can be reversed at this time. That is, are special tools or skills required to make the replacement? Can other items be damaged when the replacement is being made? The Monte Carlo Simulation must reflect the probability of events such as this.

7. Does the light bulb work after the repair has been made? If not, maybe the failure was in the switch and not in the bulb. If the bulb works now will it continue to work? In other words, is there another reason for the bulb failure? We usually assume, and rightly so, that when a light doesn't work then the bulb must be replaced. But in complex equipment, an analysis must be made of the equipment to determine the cause of the failure before the correct replacement can be made.

The question was intended to cover the check-out that is required after making a repair. In the Monte Carlo Simulation, there will be probabilities that the correct part was changed, that no damage occurred during the installation of the part and that the system checks out satisfactorily after the repair.

Tree Diagram

A tree diagram to illustrate the decision flow has been constructed and appears in Figure 1. In each branch of the tree where a "yes" or "no" appears, there is a probability corresponding to "yes" and a probability corresponding to "no", the sum of which is equal to one. Where a "G" appears, a probability density function is used to represent the possible states of nature, that is, the possible outcomes.

The probabilities and the probability density functions should be based upon historical data. In the Monte Carlo Simulation, we shall assume that the density functions are exponential unless stated otherwise and that the outcomes restricted to "yes" or "no" are Bernouli processes.

Flow Chart

If we were writing a computer program to simulate the process described, we would find a flow chart to be very helpful. The flow chart outlines the logical sequence of events to be performed by the computer. It provides us with a path from the beginning to the end of the computer run. It tells us exactly what to do, e.g., when to generate random numbers, what records to keep and when to stop.

Flow charts are usually more complicated than this one which is purposely kept simple to illustrate the concept. See Figure 2.

Example 2

Let us now turn our attention to a system that consists of ten electronic equipments identified as #1, #2, ------- #10. This equipment is contained in one location and performs the same function. For example, they could be data banks, or electronic test equipments, or something similar. Each equipment has the same failure modes and we shall identify them as A, B, C, D, and E so that there are a total of five.

Suppose that we are interested in estimating the life cycle costs for one year for equipment with an MTBF of 100 hours. We shall also assume that the model is so complex it must be simulated to estimate these costs.

From past history, we know that failures are exponential and that repair times fit a log-normal distribution. Each equipment operates for about 20 hours per day so the 10 equipments generate a total of 200 hours daily. For a 100 hour MTBF the expected number of failures is two per day. The average repair time is eight hours per failure which means that we can expect about 16 hours of repair time per day. If there are two repairmen on duty, they can work on some jobs at the same time and reduce the average repair time to six hours. But when a second job comes in, each man works on a separate job. Equipments #1 and #2 have priority. When either of them is down, both men stop what they are doing and work together to get the equipment back in working condition. Each man has a portable test rig that was provided at a cost of $2,000 each to use on a job but they do not have a back-up rig. If a rig is down for repairs they work on jobs together. There is five percent chance that a test rig will be down for repairs and the average repair time is two days.

To maintain as much simplicity as possible let us assume that there is one replacement part used for each of the five failure modes. These parts are labeled a, b, c, d, and e. During development, the test results indicated failure rates for these parts as follows:

Part	Cost	Failure Rate Per Hour	Expected Failures: Per Day	Per Wk	Per Mo
a	$ 5	.003	.6	4.2	18.2
b	10	.001	.2	1.4	6.1
c	8	.001	.2	1.4	6.1
d	20	.002	.4	2.8	12.1
e	10	.003	.6	4.2	18.2
			2.0	14.0	60.7

Based on this information, the inventory manager decides to carry an inventory corresponding to the expected demand over two months and to place on order when the inventory level reaches 9 for part a, 3 for part b, 3 for part c, 6 for part d, and 9 for part e. The average time it takes to replace a part in the inventory is normally distributed with an average of 8 days and a standard deviation of 2 days.

Students are hired during the summer to allow the repairmen to take a vacation but during the rest of the year, if either man is out, the remaining man must keep the shop going. Students can only perform simple repairs. If a wait is expected to be longer than a day, the jobs are contracted with a vendor at $200 each and will take 10 hours to repair. Hence, waiting time has a value of $20 per hour under those conditions. When equipment does not pass the check-out after repair, the failure analysis and repair will take an additional 20 hours.

With these facts available, let us list the main elements of the Monte Carlo Simulation model and then construct a flow chart. The step by step procedure is as follows:

1. Define the parameters, procedures, and policies related to operation, inventory management and maintenance.

2. Generate an equipment failure time and start the clock to keep track of the downtime.

3. Call the repairman. Does the job have a priority? Is the repairman busy or idle?

4. Is the test rig operating?

5. Is the spare part available?

6. Generate an equipment repair time.

7. Does the equipment pass the check-out after repair?

For each of these questions, we must generate a random number to determine the outcome of the situation. If there is a problem in any one of them, then other alternatives must be considered and the waiting time recorded. We shall also keep track of the cost of spares, maintenance time that cost $10 an hour for direct labor, repair and depreciation to the test rigs. The cost of maintaining an inventory amounts to .09 of the average inventory dollar and downtime is $10 per hour.

Figure 3 is a flow chart of the process indicating the decision points and the logic of the program.

We must not overlook our purpose in building such a model, which is to arrive at an estimate of the cost of operating the system over its useful life.

Probability Models

Once the flow chart has been constructed, we can design the probability models. They should correspond to the real world as much as possible to make the situation as realistic as possible which means that they should be based on historical data or historical facts. Historical data implies that an operation has been observed over a period of time and a record kept of the success and failure of the operation. From this record, a probability statement can be made.

For example, if a test rig was called upon 1,000 times during the past year and found in usable condition 990 times and not usable 10 times, then the probability of its being usable at any one time is 990/1000 or .99.

If the system is time dependent, then failure times must be recorded so that a density function can be fitted to the failure data.

Historical fact is defined here as a policy or procedure. For example, the policy may be to use only one maintenance man. As a result of this policy, we can keep track of the time when he is busy so that when failures occur we can check his status, i.e., busy or idle, which may affect waiting time and cost and produce a statistic. Thus, the historical fact leads to a quantitative result.

Taking each of the decision points in turn, let us examine in detail the logic of the Monte Carlo Simulation.

1. To generate equipment failure times we shall assume that failures are exponential and that the MTBF is 100 hours. This MTBF could be the minimum acceptable MTBF specified, or it could represent the state of the art or some negotiated value. In any case, for an MTBF of 100 hours, we can either generate the time at which the equipments fail, or take each equipment separately and for each hour of operation generate a random number, then test it to see if it corresponds to successful operation or unsuccessful operation of the equipment during that hour.

The first method uses the reliability function $R(t) = \exp(-t/MTBF)$ to generate failure times.

If the first method is used, the reliability function is written as a function of the random number and the MTBF. That is, the failure time = -MTBF x ln (random number).

This equation is derived by solving the reliability function for t, or graphically, entering the probability scale with a uniform random number and finding the corresponding time to failure. See Figure 4. Failure times are easily generated on the computer since logarithmic and uniform random number subroutines are available. The program to generate a failure time may be written as:

```
LET R = RND(-1)
LET T = -M* LOG (R)
```

Where T = failure time and M = MTBF

LOG (R) is a subroutine for computing a logarithmic

RDN (-1) is a random number subroutine

The second method requires that we check each equipment each hour to see if it fails during that hour. This would require 200 checks per day if we operate 10 equipments 20 hours per day. By selecting a small time period, we can force one of two possible decisions, either 0 failures or 1 failure. We shall use the second method to facilitate keeping track of the time.

2. Is the repairman busy? It depends upon the number

of repairmen and the status of the system. It will be necessary to identify when these men are busy and then check to see if they are busy or idle when a failure occurs.

On the computer this can be done letting a variable name equal 1 when they are busy and 0 when they are idle and then check the condition of the variable name. Hence, the probability that they are busy depends upon the MTBF, the MTTR and the number of repairmen.

3. Is the test rig in operating condition? To answer, we again generate a random number and compare it to the probability that it is working and to the probability that it is not working. If it is working, then the repair can begin; if it is not then we must have the test rig repaired and contract for the repair job if it is more than 8 hours.

4. Is the spare part available? It depends upon the inventory status. Since a running count of the inventory is maintained, a check can be made of the number on hand. If the inventory has been depleted then the job must wait for the replacement part or be sent out for contract repair.

5. Generate a repair time. The repair time is generated in the same way that a failure time was generated except for the probability function. We shall assume that repair times for a log-normal distribution are applicable. Consequently, the reliability function will differ slightly.

6. Does the equipment pass check-out? Again, we consider this as a Bernouli process and generate a random number to determine whether the equipment passes or fails check-out.

Records

While the simulation is operating a running total is maintained on failure times, repair time, waiting time, downtime, inventory levels, maintenance cost, inventory cost, cost of spares, and the depreciation on test equipment.

At the end of the run, these totals and the grand total is printed out so that we can compare the relationship between the LCC and the MTBF. It is advisable that we have as many runs as possible for each MTBF so that we can make a better decision.

Operating Procedure

To operate the model, we shall begin at time zero and generate failures then follow them through the flow chart keeping a record of the costs that occur, the inventory levels and the status of the service facilities and supporting equipment.

The cost model is essentially a matter of cumulating the costs that are experienced during the life of the equipment. The costs are those listed as life cycle cost factors and as indicated earlier they depend upon the failures generated. Time and space do not permit us to examine every failure that will be generated during the life of the equipment. But we will illustrate the procedure manually for the first few failures and summarize the results over a year's operation.

The policies were defined in the preceding section as were some of the event probabilities. Let us now define the probabilities that will be used in the simulation.

1. To generate an equipment failure, an exponential distribution with an MTBF of 100 hours is assumed for each of the 10 equipments. Hence, there is a .99 chance of no failures and a .01 chance of one failure. We shall generate a random number for each equipment for each hour, if it is .01 a failure has occurred, if it is any other number a failure has not occurred.

2. The chance of the repairman is busy depends upon the status of the system and must be checked at the time the failure occurs.

3. About 12% of the time one repairman will not be available; and about 2% of the time neither repairman will be available. A random number is generated to see if there are 0, 1 or 2 men available.

4. A test rig has a reliability of .95 but will be available only 90% of the time, therefore, both rigs will be available 81% of the time. We again must generate a random number and compare it to the parameters to determine what is available.

5. The chance of a spare part not being available depends upon the many factors affecting the inventory levels. This will be determined at the time of the failure.

6. To generate a repair time, a log-normal distribution is assumed with an average of 8 hours and a standard deviation of 2 hours.

7. The chance that the equipment passes the checkout test after repair is.98. We generate a random number. If it is .98 or less then the equipment passes, if it is .99 or .00 then it does not.

The Operation

To illustrate the operation of the model as the computer would do it, the data in Table I has been generated. This table lists the time at which the failure occurred, when the repair began, the repair time, when the repair was completed, the waiting line and other aspects of the model that have been discussed.

With the flow chart in Figure 4 as a guide, we shall step through the first part of the Monte Carlo Simulation model as though the computer were in operation except we shall perform the operations as at a much slower pace. The discussion that follows will be based upon the flow chart and Table I.

The first failure occurred during the first hour at which time module "A" failed on equipment #5. When the status of the supporting elements were tested, both repairmen and both test rigs were available, the service facility was available and the repair part was in inventory. A second random number was generated to determine the repair time which was 17 hours, so the job was expected to be completed by the end of the 19th clock hour if it began service at the start of the 2nd hour.

If we plot the simulation throught the flow chart, we see that for the first four equipments there was no failure and no jobs were completed so we are brought back to the first part of the chart to update the clock and generate another random number. When the random number is generated for equipment #5, a failure occurs and the flow chart directs us to facilities available (yes), men available (yes), rigs available (yes), part in inventory (yes). To decide if the path is through "yes" or "no", random numbers are generated and tested against the

parameters defined when the program was initialized. A random number is generated to determine the length of the repair time (17 hours), and repair begins. At this time, the system status must be updated by reducing the number of service men, number of test rigs, and number of facilities available by one, reduce the inventory for part a by one, move the equipment counter up to equipment #6, set the job completion counter to 18 and set priority status counter to "no".

The second failure occurred during the 6th hour with one man and one rig available. The repair time is 5 hours so the job should be completed at the end of the 12th hour.

On the flow chart failure #2 follows the same path as failure #1. When this job was completed (at 12th hour) it was checked and since it checked out satisfactory, the equipment was returned to service, the facility, man, and rig made available for the next failure. Downtime was recorded as was the type of failure, equipment number and the cost of the repair.

During the 13th hour, equipment #1 failed and it can go directly into the service facility because it is a priority job. Both repairmen are directed to the job. Therefore, failure #1 which had 5 hours of repair time remaining must be placed in a waiting status while failure #3 is being repaired. On the flow chart repair is halted and then the men are assigned to the task of repairing the priority job.

The repair on failure #3 with two men working together was 14 hours so the job should be completed by the 27th hour. This resulted in job #1 not being completed until the 32nd hour resulting in a waiting time of 13 hours. The computer updated the completion time for failure #1 and proceeded to repair failure #3.

Table I would be easier to follow if there were no priority jobs. When a priority repair job arrives, all other work must stop and this is why the repair times and waiting times have been changed for failure #1, 6 and so on. To illustrate the effects of a priority job, a time-event chart has been constructed for the first 70 hours of operation and can be seen in Figure 5. This figure shows that during the first 67 hours of operation there were 61 hours of waiting time. It also shows that 5 equipments were out of service at one point in time and that the idle time was about 8 hours. The graph in Figure 5 illustrates one of the advantages of a simulation, that is, it allows us to "see" how a system reacts to a given set of parameters.

If you were to continue to step through the model using Table I and Figure 4 you will arrive at the completion times and waiting times as indicated. The flow chart does not describe the procedure for updating the system or collecting costs since that is a computer programming problem. Some of the reference material describe the procedure in great detail.

The Analysis

In Table I the results obtained when the system is operated for 450 hours are listed. For the 34 failures 361 hours of waiting time were accumulated for an average of 10.6 hours of waiting time per failure. The repair time was surprisingly close to the waiting time, 397 hours, for an average repair time of about 11.7 hours. We see also that on 3 occasions the equipment did not pass check-out, that test rigs were not available on a few occasions and that repairmen and/or facilities were lacking at various times. But the biggest factor causing the increase in waiting time are the 6 priority repair jobs. Also, the life cycle costs are being collected in the program but do not appear in Table I. They are:

Maintenance Cost	397 hours @ $10/hour=$	3,970
Spares Cost	34 spares (per table)	379
Downtime & Waiting Time	397 @ $10 & 361 @ $20=	11,190
Test equipment repair	3 breakdown @ $100/ea=	300
Contract repair	3 contracts @ $200/ea=	600
Inventory Cost	9% of average inventory=	96
		$16,535

It is readily apparent that about two-thirds of the LCC collected so far are attributed to downtime and waiting time. The important point to remember is that they are dependent upon the MTBF and MTTR and by running a Monte Carlo Simulation using different combinations of the MTBF and MTTR we can derive estimates of these costs as well as other costs that occur during the life cycle for each MTBF/MTTR combination.

A run of 450 hours is not sufficient for most problems, nor is a single run sufficient. In Table II, the results of 20 runs for an MTBF of 100 hours and an MTTR of 8 hours are listed.

The table does not contain all the details that are in the computer program. However, it does illustrate the random nature of the simulation and the range of expected values. For example, the total number of failures ranges from a low of 445 to a high of 571 and average out to 507 as compared to an expected number of failures of 500. From this range we can judge the relative accuracy of an estimate.

The model was carried one step further by examining other combinations of the MTBF and MTTR in a simulation run. The MTBF of 100 hours and 200 hours were examined in all combinations with an MTTR of 6 hours and 8 hours. The results are contained in Table III where point estimates have been listed for the various costs.

The acquistion price is not listed but we know that the acquisition price is usually a function of the MTBF and MTTR. If the acquisition price were approximately equal to 100 x MTBF then option D would be the optimum when all costs are considered. However, if the acquisition price were approximately equal to $(MTBF)^2$ x 100/MTTR then option C would be the optimum. In any case, the point is that the optimum solution depends upon the total cost, i.e., acquisition price plus LCC; and Monte Carlo Simulation provides us with a means of estimating the optimum.

Conclusion

In this paper, it was the author's intention to illustrate how a Monte Carlo (or stochastic) Simulation model could be designed to estimate the life cycle costs of an equipment. The main advantage to this type of approach is that it provides the designer with a tool that allows him to examine various design alternatives prior to production. The user can also examine various policies and procedures to determine their affect on system cost.

The steps that should be considered can be summarized as follows:

1. Define the problem.

2. Define the factors related to the problem and the parameters under study.

3. Build a mathematical model and test the parameters.

4. Write a computer program.

5. Select the experimental design that will provide the data for solving the problem.

6. Run the program and analyze the output.

Except for step 4, this is what we have done. We viewed the problem through the eyes of the designer, i.e., what MTBF and MTTR should be used?

Several life cycle cost factors were discussed and some of them were included in the examples presented. All of them can be included in a simulation model and for this reason the Monte Carlo Simulation approach is superior to the analytical approach which may be too difficult to model in such cases.

A Monte Carlo Simulation math model uses probability density functions to generate events and realistically portray the interacting functions that take place between men and machines. Some of the models and generators were discussed and used in the example.

The design consisted of 2 factors, MTBF and MTTR, at two levels each. Any number could have been included at many more levels but neither time nor space permitted us to do so.

When the runs were completed, we could see the tradeoffs that were possible between acquisition price, MTBF, MTTR and LCC before the design was finalized.

At the present time the model is being used in a realiability simulation exercise to help the contractors optimize their designs. In real life, models are being developed for use so that the parties in a contract can make better use of their funds.

In conclusion, models such as this can be used successfully, and also, provide the use with a more realistic estimate of life cycle costs. They provide a better view of the affects that policies and procedures have upon the system and its operation.

Bibliography

Barton, Richard F., A Primer on Simulation and Training, Prentice-Hall Inc., Englewood Cliffs, N. J., 1970

Forrester, Jay W., Industrial Dynamics, The MIT Press and John Wiley and Sons, 1961

Fulmer, John L., Business Simulation Games, South-Western Publishing Co., Cincinnati, Ohio, 1963

Gordon, Geoffrey, System Simulation, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1969

McKenney, James L., Simulation Gaming for Management Development, Harvard Business School, Boston, Mass., 1967

Meier, Robert C., Newell, W. T., and Pazer, H. L., Simulation in Business and Economics, Prentice-Hall Inc., Englewood Cliffs, N. J., 1969

Mize, Joe H. and Cox, J. G., Essentials of Simulation, Prentice-Hall Inc., Englewood Cliffs, N. J., 1968

Naylor, Thomas H., Balintfy, J. L. and Burdick, D.S., Computer Simulation Techniques, John Wiley and Sons, New York, N. Y., 1968

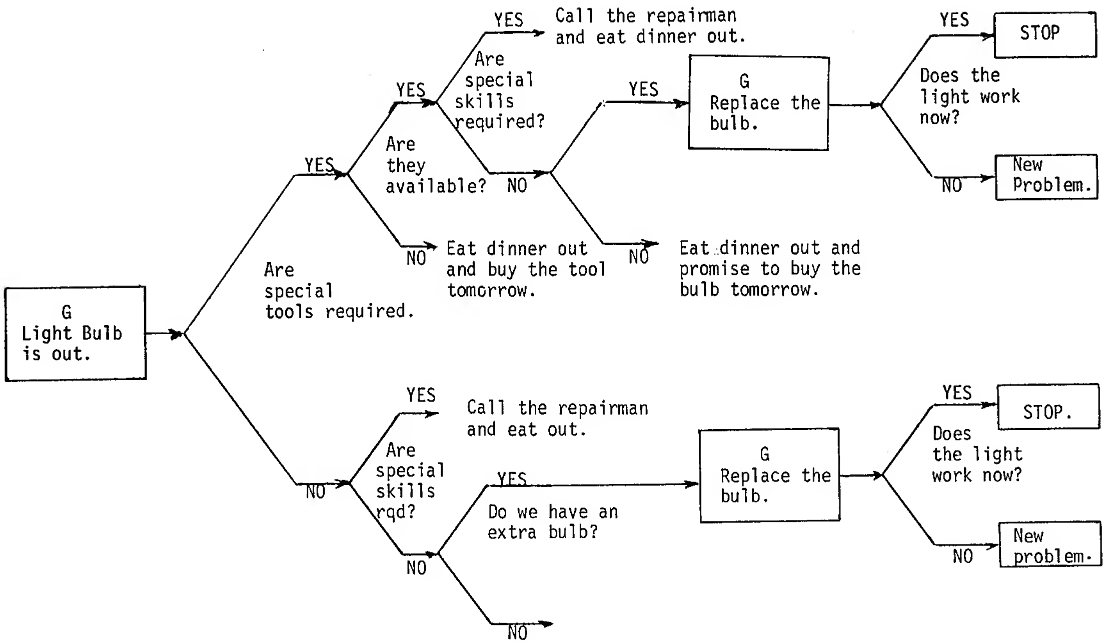


Figure 1. Tree Diagram of the Light Bulb Problem.

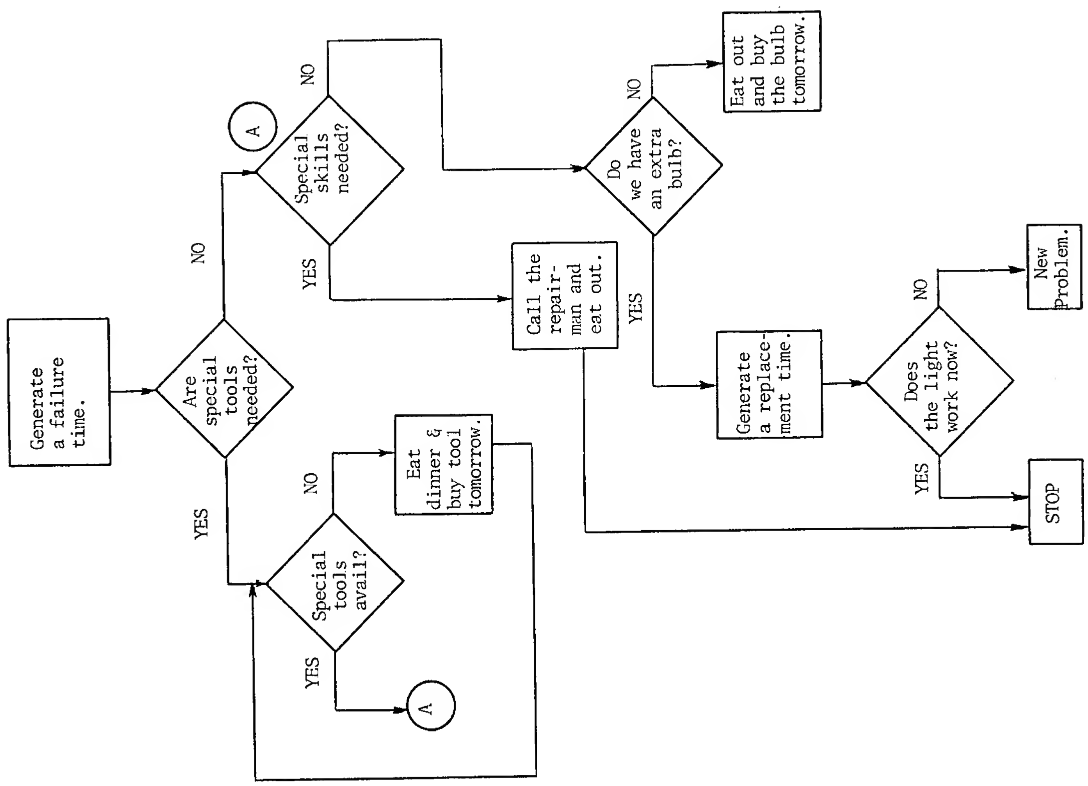


Figure 2. Flow Chart for the Light Bulb Problem.

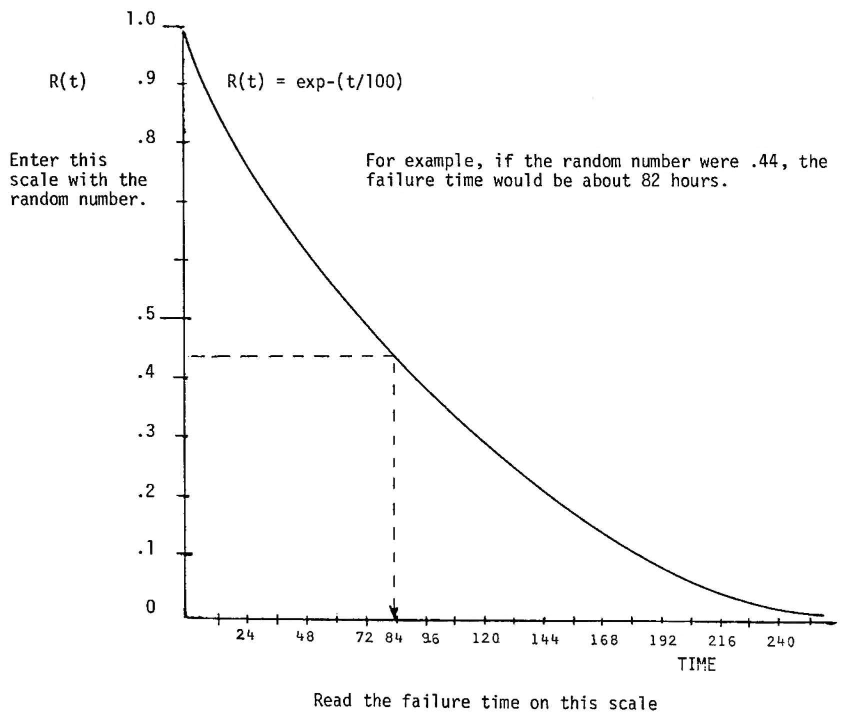


Figure 3. Graphical Illustration of Failure Time Generation.

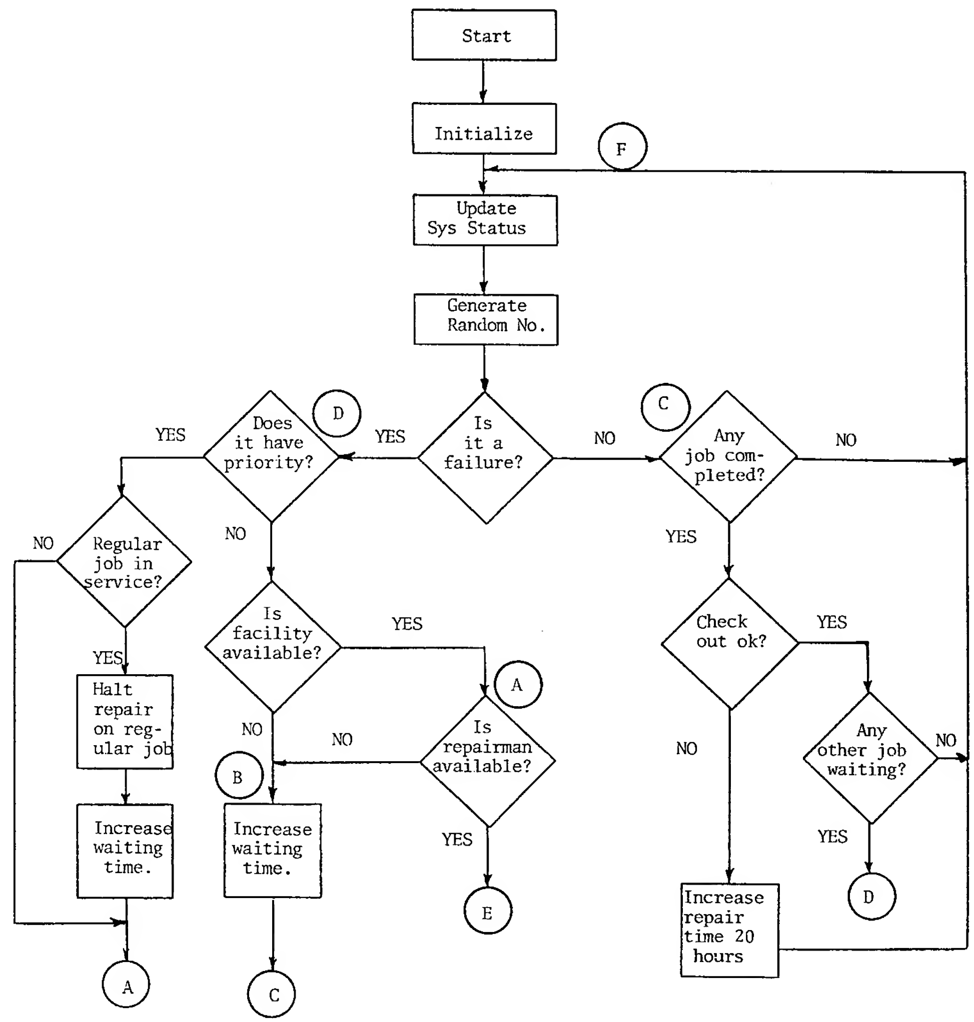


Figure 4. Flow Chart for Example 2.

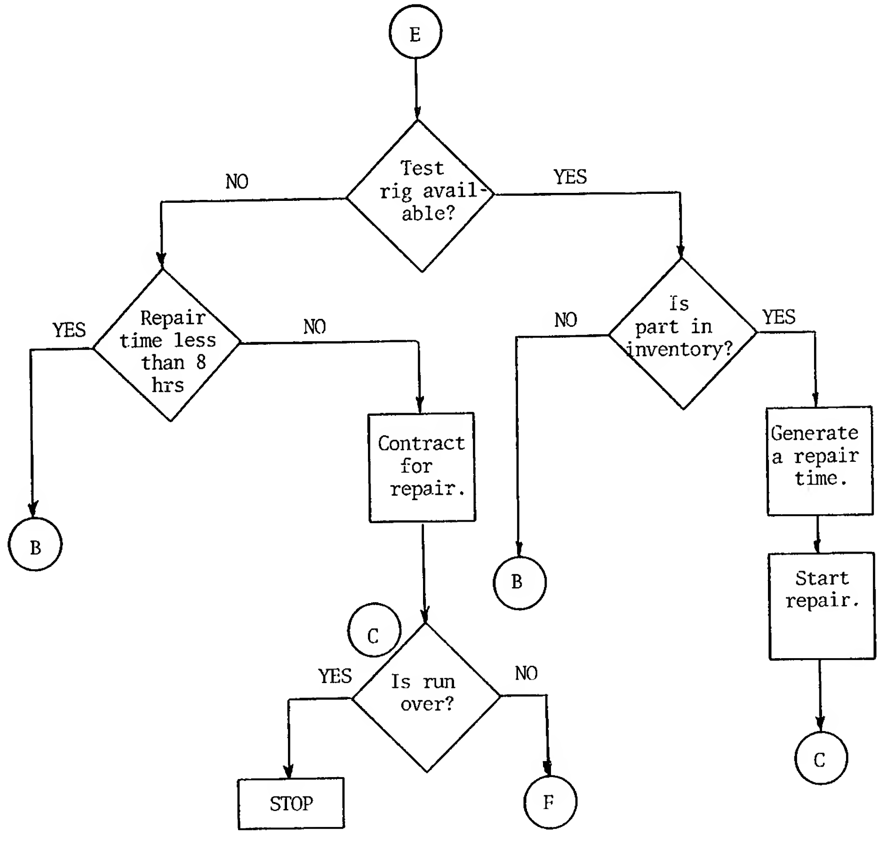


Figure 4. Continuation of Flow Chart for Example 2.

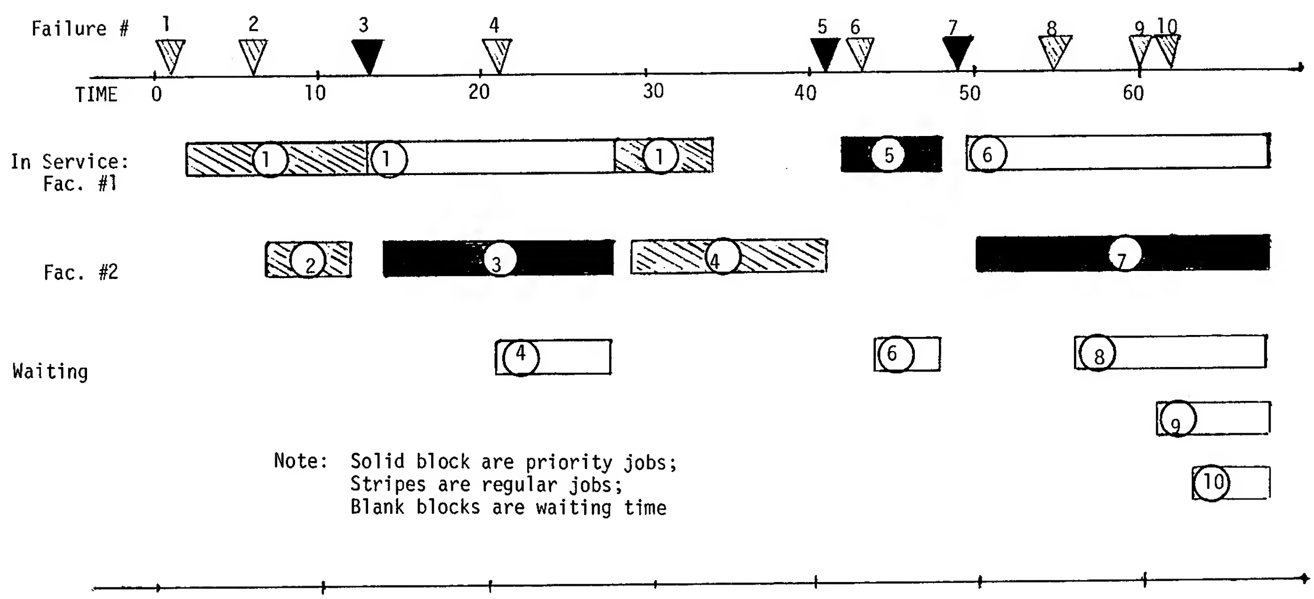


Figure 5. Time-Event Chart for the First 70 Hours of the Monte Carlo Simulation.

TABLE I

DATA COLLECTED DURING THE FIRST 450 HOURS OF THE MONTE CARLO SIMULATION

Failure Number	Module Failed	Failure Time	Begin Repair	Repair Time	Repair Completed	Pass Checkout?	Waiting Time	No. of Facilities Available	No. of Repairmen Available	Test Rigs Available	No. of Equip out of Svc.
1	A	1	2	17	1~~9~~ 34		13	2	2	2	1
2	E	6	7	5	12			1	1	1	2
3	C	13	14	14	28			0	0	0	2
4	D	21	29	12	41		8	1	0	1	3
5	E	41	42	6	48			1	1	2	2
6	D	43	4~~9~~ 68	4	5~~3~~ 72		~~6~~ 23	1	1	1	2
7	C	49	50	17	67			1	1	1	2
8	D	55	68	25	93		13	0	0	0	3
9	A	60	73	6	79		13	0	0	0	4
10	E	62	80	12	92		18	0	0	0	5
11	B	72	93	3	96		21	0	0	0	6
12	A	110	111	3	114			2	2	2	1
13	A	121	122	6	128			2	2	2	1
14	A	140	141	13	154			1	2	1	1
15	A	150	155	3	158		5	0	1	1	2
16	E	159	160	~~8~~ 28	1~~6~~8 198	No		1	2	1	1
17	E	199	200	8	208			1	2	1	1

TABLE I (Cont'd)

Failure Number	Module Failed	Failure Time	Begin Repair	Repair Time	Repair Completed	Pass Checkout?	Waiting Time	No. of Facilities Available	No. of Repairmen Available	Test Rigs Available	No. of Equip out of Svc.
18	D	247	248	10	258			2	2	2	1
19	C	255	256	6	262			1	1	1	2
20	E	269	270	8	278			2	2	2	1
21	E	274	~~279~~ 308	~~17~~ 31	~~290~~ 357	No	~~34~~ 53	1	1	1	2
22	E	277	279	28	307		2	0	0	0	3
23	E	278	308	4	312		30	0	0	0	4
24	A	280	313	7	320		33	0	0	0	4
25	D	300	~~321~~ 358	6	~~327~~ 364		~~27~~ 57	0	0	0	5
26	D	320	~~328~~ 353	10	~~338~~ 363		~~8~~ 23	0	0	0	3
27	E	321	322	30	352			0	0	0	3
28	A	321	364	3	367		43	0	0	0	4
29	B	358	365	8	373		7	0	0	0	4
30	A	375	376	13	~~389~~ 409	No		2	2	2	1
31	E	386	387	13	400			1	1	1	2
32	D	411	412	6	418			2	1	2	1
33	D	418	419	30	449			2	2	2	2
34	D	435	436	2	438			1	1	1	2
TOTALS				397			361				

TABLE II

DATA COLLECTED DURING 20 RUNS FOR AN MTBF OF 100 HOURS AND AN MTTR OF 8 HOURS

RUN	A	B	MODULE C	D	E	TOTAL # OF FAILURE	SPARES COSTS	FAILURE HOURS
1	143	40	38	98	158	477	$4,959	3823
2	138	45	44	95	185	507	5,242	4067
3	165	46	54	116	140	571	5,437	4574
4	140	53	49	104	145	491	5,557	3938
5	126	47	53	92	141	459	4,774	3672
6	142	61	50	84	183	520	5,294	4160
7	147	56	43	108	178	532	5,579	4256
8	128	45	57	118	151	499	5,416	3992
9	154	44	54	107	149	528	5,272	4224
10	160	62	58	114	147	541	5,634	4328
11	143	45	42	109	150	489	5,181	3912
12	126	50	55	117	156	454	5,470	3632
13	159	62	64	97	185	567	5,717	4536
14	142	63	54	93	156	508	5,192	4064
15	152	70	47	81	155	505	5,006	4040
16	131	41	54	96	170	492	5,117	3936
17	127	44	39	93	142	445	4,667	3560
18	154	53	43	97	146	493	5,044	3944
19	162	54	55	108	158	537	5,530	4296
20	155	55	46	95	182	533	4,913	4264

TABLE III

SUMMARY OF 20 RUNS OF THE MONTE CARLO SIMULATION WITH DIFFERENT VALUES FOR THE MTBF AND MTTR

		A	B	C	D
	OPTION	A	B	C	D
	MTBF	100	200	100	200
	MTTR	8	8	6	6
COST :					
	Maintenance	$40,560	$19,712	$39,760	$15,222
	Spares	5,250	2,284	5,170	2,621
	Downtime	61,700	20,530	40,960	7,420
	Waiting Time	135,883	43,190	91,421	31,419
	Test Equipment	5,700	5,300	4,700	6,300
	Contract Repair	3,900	2,100	4,200	1,500
	Inventory Cost	1,460	697	1,376	728
Total		$254,453	$93,813	$187,587	$65,210

MODELING THE BATHTUB CURVE

Thomas W. Calvin

IBM Components Division
East Fishkill Facility
Hopewell Junction, New York 12533

Attempts to model a "bathtub" curve often assume that useful life (the flat portion of the curve), and wear-out (the increasing portion) begin after time zero. The resulting models are called either composite or mixed, depending on the method of model building. Infant mortality, random, and wearout failure modes are relegated to the appropriate portions of the bathtub curve. An approach is suggested in which all models of failure begin at time zero. By looking at the unit (device, component, machine, etc.) as a serial system with respect to failure mechanisms in which any failure results in unit failure, reliability is the product of the reliabilities of the individual mechanisms; equivalently the cumulative hazard is the sum of the individual cumulative hazards.

Introduction

During the lifetime of equipment or parts, many situations result in equipment breakdown. Early in this lifetime, catastrophic or sudden failures occur due to initial existing weakness or defects. As these failures are replaced, defective product is removed, producing a decreasing failure rate. Under continued operation, parts begin to deteriorate producing wear-out or delayed failures. Failures which did not exist initially are being generated causing an increasing failure rate. Some failures occur when design strengths are exceeded by environmental stresses which are frequently difficult to determine.

This combination of failures produces the familar bathtub failure rate curve (Fig. 1) often associated with part or equipment performance. Early failures first dominate, exhibiting the decreasing failure rate characteristic of the infant mortality period. Finally deterioration failures dominate with the increasing failure rate indicative of wearout. Between these two extremes is an essentially flat region (when viewed on linear coordinates) where neither type failure domin-

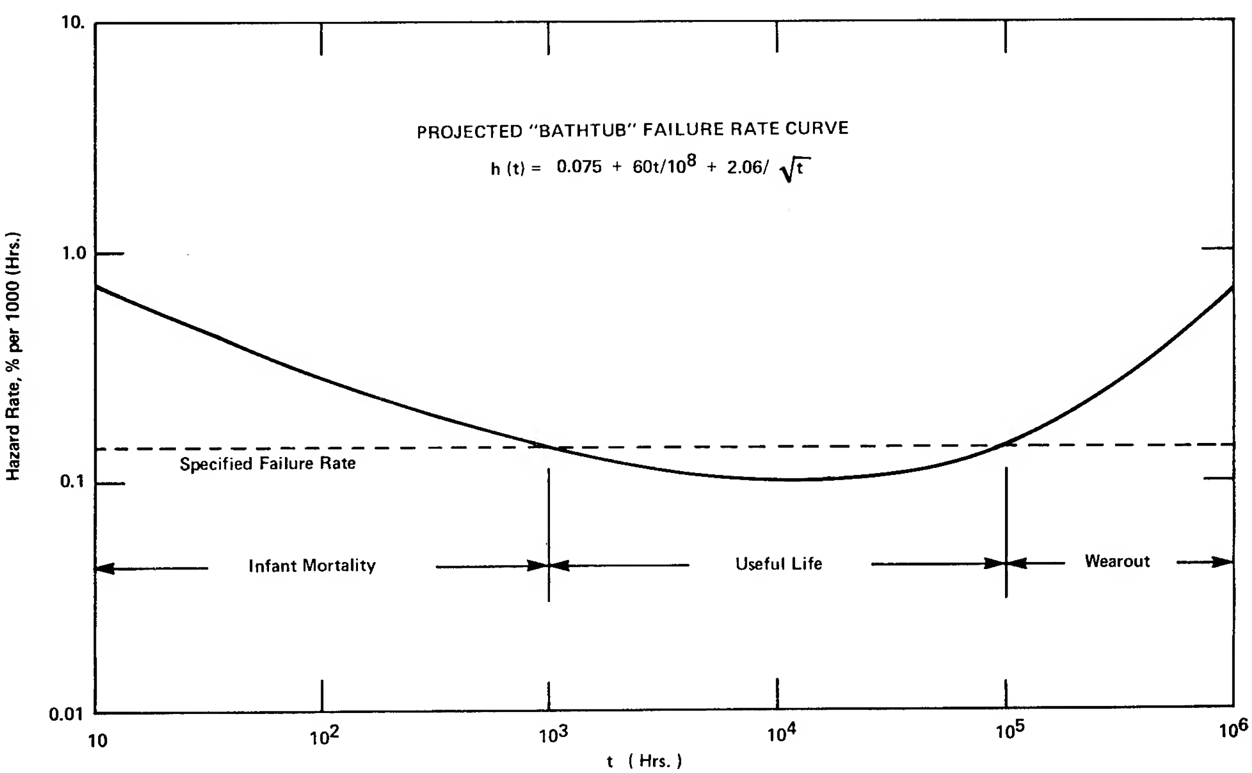


Fig. 1. Bathtub curve.

ates. Most failures in this region appear to occur at random due to many unrelated, unpredictable causes. This is the useful life period within which the failure rate is generally specified for the equipment or part.

A number of models have been suggested to fit this curve. With many failure mechanisms present, some of these models become complex and difficult to estimate. This paper proposes an approach that can accomodate any number of failure mechanisms conveniently. It is necessary, however, to isolate, identify, and evaluate these mechanisms separately. The bathtub model is then "built" from these individual mechanisms.

Some Definitions and Formulas

Failure Modes and Mechanisms

Failure mechanism implies the type of failure, such as migration, diffusion, cracked substrates, corrosion, electrical overloads, leaking seals, etc. Depending on their nature, these mechanisms can occur early, late, or throughout the life cycle. Some mechanisms result from manufacturing defects; others are caused by continuous operation. Defective product tends to fail early and deteriorated product much later in the life cycle.

Failure mode refers to the location of the failure in the life cycle. Early failures due to initial weaknesses or defects that have escaped detection are in the infant mortality mode. These escapes can be due to screening inefficiency or inability to detect certain defect types. Such failures are incapable of continued performance at operating conditions. These include time zero failures awaiting detection, "weak sisters" which fail once full operating conditions are attained, and intermittent failures which alternately perform and fail depending on the length of time at operating conditions.

As soon as an equipment or part is in operation, deterioration can begin. Failures resulting from this deterioration constitute the wearout mode. Both perfect and imperfect product can contribute to this mode. Perfect product can fail due to overstressing as a result of marginal design or improper application. Imperfect product is "good" in that it meets the manufacturing specifications and can give adequate performance. State-of-the-art and economic limitations often restrict the improvement of specifications, making it impossible to produce perfect product consistently.

Finally, there is the random mode of failure resulting from a variety of causes often unrelatable to time. Such failures occur throughout product lifetime and are responsible for a large portion of the useful life period in which infant mortality and wearout are minimal. These random occurrences, often related to environmental stresses or upsets, would include power shutdowns, power surges, mechanical damage, etc. Many minor failures whose occurrence is too seldom to be treated separately can appear random when considered as a group.

Weibull Distribution

The Weibull cumulative distribution is defined as

$$F(t) = 1 - \mathrm{EXP}[-(t-\gamma)^m/\alpha^m] \qquad (1)$$

for $t \geq \gamma$; α, m positive

α = scale parameter
m = shape parameter
γ = location parameter

The first derivative of F(t) with respect to t is the Weibull density function; thus

$$dF(t)/dt = f(t) = \frac{m(t-\gamma)^{m-1}}{\alpha^m}\ \mathrm{EXP}\ \frac{-(t-\gamma)^m}{\alpha^m} \qquad (2)$$

Reliability Function

$$R(t) = 1 - F(t) \qquad (3)$$

which for the Weibull distribution is

$$R(t) = \mathrm{EXP}[-(t-\gamma)^m/\alpha^m] \qquad (4)$$

Hazard or Instantaneous Failure Rate

The hazard or instantaneous failure rate is defined as

$$h(t) = f(t)/R(t) \qquad (5)$$

resulting in a Weibull hazard of

$$h(t) = m(t-\gamma)^{m-1}/\alpha^m \qquad (6)$$

Cumulative Hazard

Integrating h(t) dt over t yields the cumulative hazard, H(t)

$$h(t)dt = H(t) = -\ln[1 - F(t)] \qquad (7)$$

or,

$$H(t) = -\ln R(t); \qquad (8)$$

$$F(t) \approx H(t) \text{ for } F(t) < 0.1$$

The cumulative Weibull hazard is

$$H(t) = (t-\gamma)^m/\alpha^m \qquad (9)$$

Proposed Bathtub Model

Prior Models

Some of the proposals existing in the literature for fitting bathtub curves will be discussed briefly. One method is to divide the curve into three regions: infant mortality, useful life, and wearout. Each region is approximated by a straight line resulting in a piecewise-linear fit. Accuracy can be improved by taking more segments, striking a balance between goodness-of-fit and computational complexity.

A variation of this approach is a composite model where each segment is a section of a probability dis-

tribution with a non-zero location parameter. The disadvantage of these two approaches is that the model is described by a series of separate equations or distributions; a number of partition parameters are necessary to define each segment location.

A better approach is the mixed distribution function:

$$F(t) = \sum_{i=1}^{k} P_i F_i(t) \text{ where } 0 \le P_i \le 1 \quad (10)$$

$$\text{and } \Sigma P_i = 1$$

$F_i(t)$ is the ith sub-population; the quantities P_i, called mix parameters, are proportions of the subpopulation mix.

$$R(t) = \sum_{i=1}^{k} P_i R_i(t) \quad (11)$$

$$\text{and } H(t) = \ln R(t) = - \ln \left[\sum_{i=1}^{k} P_i R_i(t) \right] \quad (12)$$

This model could accomodate many failure sub-populations, but has the disadvantage of requiring many mix parameters to describe the sub-population mix. The hazard function is moderately complex; this approach would be difficult to apply for many sub-populations. It is very convenient and easy to use with only two sub-populations.

Proposed Model

The model presented in this paper produces one continuous equation requiring the estimation of less parameters. The hazard function is a simple additive model which is easy to evaluate by considering the failure rates for each mechanism separately. A model in which each mechanism requires only a scale and shape parameter is assumed adequate for most situations but location parameters can be added where necessary.

With the assumption that all failure mechanisms and hence modes begin at time zero, the failure modes (infant mortality, wearout, and random) compete throughout product life. Some mechanisms dominate the early failures while others dominate at the end of useful life, though both appear to a lesser extent during the useful life period. The decreasing infant mortality overlaps the increasing wearout to flatten the curve. This flat portion is further enhanced by random failures which largely contribute during this period. Thus, the failure rate during any time interval is a sum potentially including all modes of failure. With this concept of failures and a Weibull distribution assumption for each mechanism, an adequate, relatively simple bathtub model can be developed which supports the intuitive addition of individual mechanism failure rates to obtain the total equipment or part failure rate.

Assuming each failure mechanism and mode is independent, equipment or parts can be viewed as a serial system in which any failure results in unit failure. Under these conditions, the reliability is the product of individual reliabilities:

$$R(t) = \prod_{j=1}^{k} R_j(t) \quad (13)$$

The cumulative hazard is the sum of the individual hazards:

$$H(t) = \sum_{j=1}^{k} H_j(t) \quad (14)$$

where k is the number of failure mechanisms and modes. Assuming each mechanism is Weibull distributed beginning at time zero, the cumulative hazard, from eq. (9), is

$$H(t) = \sum_{j=1}^{k} (t/\alpha_j)^{m_j} \quad (15)$$

It is convenient to separate the infant mortality, random, and wearout modes, yielding

$$H(t) = \Sigma (t/\alpha_i)^{m_i} + \Sigma (t/\alpha_r) + \Sigma (t/\alpha_w)^{m_w} \quad (16)$$

in which each mode above (identified by i, r, and w respectively) contains any number of mechanisms. To simplify the remaining formulas, assume one failure mechanism per mode:

$$H(t) = (t/\alpha_i)^{m_i} + t/\alpha_r + (t/\alpha_w)^{m_w} \quad (17)$$

The following models are readily generalized to many mechanisms by adding Σ or Π as appropriate.

The reliability function, from (9) and (13),

$$R(t) = [\mathrm{EXP} - (t/\alpha_i)^{m_i}] \, [\mathrm{EXP} - (t/\alpha_r)]$$

$$[\mathrm{EXP} - (t/\alpha_w)^{m_w}] \quad (18)$$

The bathtub curve (hazard function)

$$h(t) = m_i t^{m_i - 1} / \alpha_i^{m_i} + 1/\alpha_r + m_w t^{m_w - 1} / \alpha_w^{m_w} \quad (19)$$

The density function, from (5),

$$f(t) = h(t) \, R(t)$$

$$f(t) = (m_i t^{m_i - 1} / \alpha_i^{m_i} + 1/\alpha_r + m_w t^{m_w - 1} / \alpha_w^{m_w})$$

$$[\mathrm{EXP} - (t/\alpha_i)^{m_i}] \, [\mathrm{EXP} - (t/\alpha_r)] \, [\mathrm{EXP} - t/\alpha_w)^{m_w}] \quad (20)$$

This approach produces relatively simple additive hazard or failure rate models. A more useful form of these models can be obtained by replacing the scale parameters (α) by specific cumulative hazard rates at a particular time as follows:

$$H_i = (t_i/\alpha_i)^{m_i}$$

$$H_w = (t_w/\alpha_w)^{m_w}$$

$$H_r = t_r/\alpha_r$$

thus

$$\alpha_i^{m_i} = (t_i^{m_i}/H_i)$$

$$\alpha_w^{m_w} = (t_w^{m_w}/H_w)$$

$$\alpha_r = t_r/H_r$$

Substituting into the cumulative hazard model (17)

$$H(t) = H_i (t/t_i)^{m_i} + H_r (t/t_r) + H_w (t/t_w)^{m_w} \qquad (21)$$

Similarly for the hazard or instantaneous failure rate:

$$h_i = (m_i/t_i)\,(t_i/\alpha_i)^{m_i}$$

$$h_w = (m_w/t_w)\,(t_w/\alpha_w)^{m_w}$$

$$h_r = 1/\alpha_r$$

and

$$\alpha_i^{m_i} = m_i\, t_i^{m_i - 1}/h_i$$

$$\alpha_w^{m_w} = m_w\, t_w^{m_w - 1}/h_w$$

$$\alpha_r = 1/h_r$$

then

$$h(t) = h_i (t/t_i)^{m_i - 1} + h_r + h_w (t/t_w)^{m_w - 1} \qquad (22)$$

also

$$h(t) = m_i H_i\, t^{m_i - 1}/t_i^{m_i} + H_r/t_r + m_w H_w\, t^{m_w - 1}/t_w^{m_w} \qquad (23)$$

Eqs. 22 and 23 are useful for evaluating data graphically and for building a projection model based on specified or evaluated failure rates. The following examples illustrate the convenience of using these models.

Model Application Examples

Graphic Analysis of Simulated Data

Over 10^4 hours, 49 components have failed out of 10^5 components. Analysis of these failures revealed four failure mechanisms (designated A, B, C and D) whose times to failure are listed in Table I. Using a Weibull distribution within mechanisms, the times to failure were simulated assuming the percent random error in the time to failure followed a normal distribution.

I. Simulated data: times to failure

Failure Mechanisms			
A	B	C	D
504	118	413	4310
990	390	1561	6070
1540	840	3540	8400
1908	1609	6310	8750
2640	2385		
2979	3220		
3500	3510		
3820	4570		
4530	5060		
4940	5590		
5590	6190		
5640	6630		
6730	7190		
7310	8060		
7540	8310		
7960	8440		
8180	9140		
8800	9375		
9125	9480		
9360	9950		
9740			

A cumulative hazard log-log plot was used to estimate the parameters for the bathtub model, since from (9):

$$H(t) = (t/\alpha)^m \quad \text{and}$$

$$\log H(t) = m \log t - m \log \alpha$$

For this example, the cumulative hazard approximately equals the cumulative fraction failed:

$$10^5 H(t) \approx 10^5 F(t) = 10^5\, r/10^5 = r$$

where r represents the total failures at t and the 10^5 multiplier converts the failure rates to % per 1000 hrs.

All times to failure are first combined and plotted, then separated into the individual mechanisms. The log-log plot of the combined data (Fig. 2) is a reason-

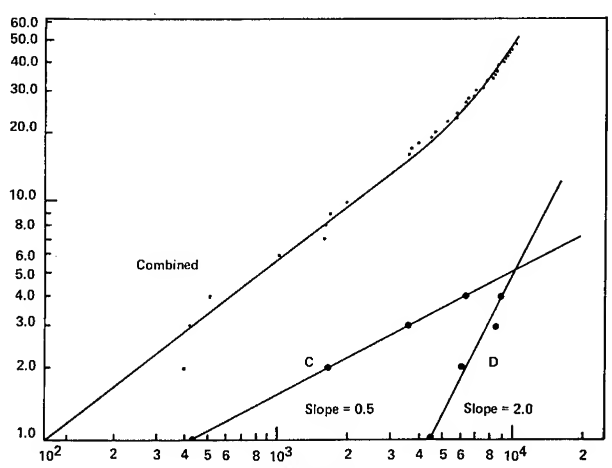


Fig. 2. Simulated data.

ably straight line with 0.75 slope to 3500 hours, where it begins to curve upward. Thus a Weibull assumption due to this departure from linearity does not adequately describe the data.

When the data is separated into individual mechanisms, only B (Fig. 3) appears to depart from linearity. A (Fig. 3), C and D (Fig. 2) are reasonably

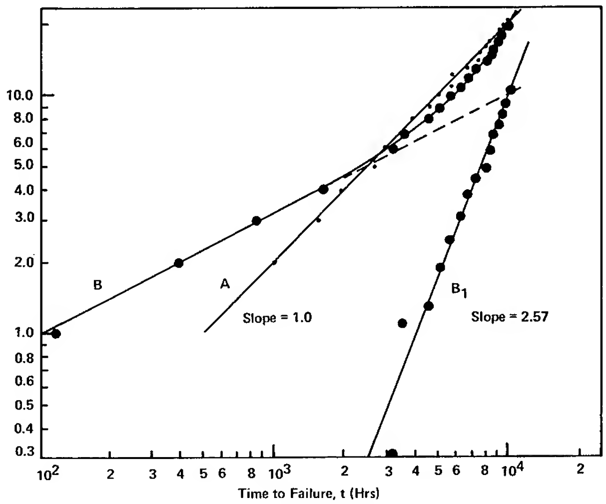


Fig. 3. Simulated data.

straight-line fits with Weibull shape parameters or slopes of 1.0 (random), 0.5 (infant mortality), and 2.0 (wearout), respectively. Mechanism B is linear up to about 2500 hours with an 0.5 slope, curving upward thereafter. Considering this mechanism to be made up of two modes of failure, these modes can be separated as indicated in Fig. 3. Fit a straight line to the data up to 2500 hours and extend this line to 10,000 hours (dashed line). For each t greater than 2500 hours, subtract the cumulative hazard on the line from the actual cumulative hazard. These differences plot against t as a straight line with 2.57 slope. Line B_1 represents the wearout mode of mechanism B while the extended (dashed) line depicts the infant mortality mode. Sufficient information is available to fit the bathtub curve. From (23),

$$h(t) = \sum_{j=1}^{k} m_j H_j t^{m_j - 1} / t_j^{m_j}$$

where k is the number of mechanisms and modes. From the graphs the m_j's are the slopes of the lines and the H_j's are obtained for particular t_j's. Using $t_j = 10^4$ the equation is (in % per 1000 hours):

$$h(t) = .5(5)t^{-5}/(10^4)^{.5} + 2(5)t/(10^4)^2 + 20/10^4 +$$

$$.5(10)t^{-.5}/(10^4)^{.5} + 2.57(10.5)t^{1.57}/(10^4)^{2.57}$$

Simplifying the bathtub curve,

$$h(t) = .002 + .075/t^{.5} + t/10^7 + 1.415t^{1.57}/10^9$$

(Fig. 4)

Projecting A Failure Rate Curve

A new level of integration in semiconductor technology is ready to be introduced. It is desired to project the component lifetime up to 100K hours. After a comprehensive investigation, the following failure sources were determined and classified as random, infant mortality, and wearout.

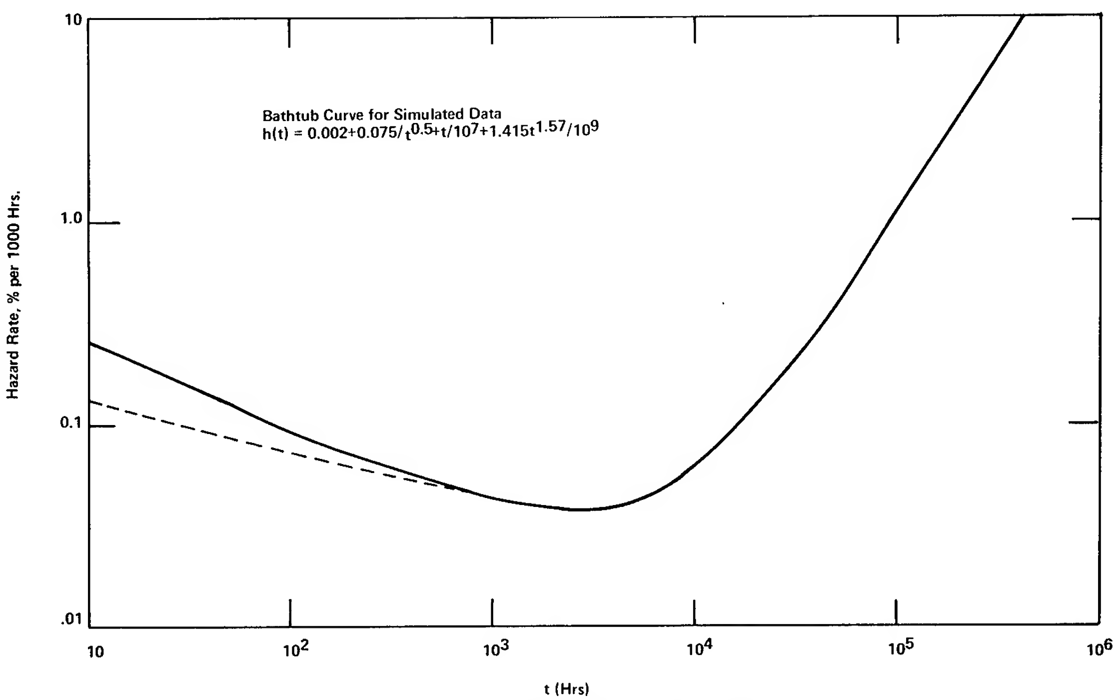


Fig. 4. Simulated data bathtub curve.

Mechanism Source	Mode	Failure Rate
Semiconductor	r	0.015%/k hrs
Metallurgy	w	0.040%/k hrs at 100k hr
Insulation	r	0.060%/k hrs at 100k hr
Interconnection	i	0.030%/k hrs at k hr
Corrosion	w	0.020%/k hrs at 100k hr
Substrate	i	0.025%/k hrs at k hr
Package	i	0.010%/k hrs at k hr

In addition, it was discovered that all of the infant mortality failures had a Weibull shape parameter, $m = 0.5$, and $m = 2.0$ was adequate for wearout. Hence, the constants required for the bathtub curve were known; the resulting equation is (in % per k hrs):

$$h(t) = (0.015 + 0.060) + \frac{0.040 + 0.020}{100,000} t + (0.030 + 0.025 + 0.010)\sqrt{\frac{1000}{t}}$$

$$h(t) = 0.075 + 60t/10^8 + 2.06/\sqrt{t} \qquad \text{(Fig. 1)}$$

To monitor performance in the field, any failure rate appearing above the curve would suggest the new technology was not meeting its projection. The specified failure rate for this technology is 0.14% per 1000 hours (dashed line, Fig. 4) achieved between 1000 and 10^5 hours. When severe wearout commences beyond the region of interest (here 10^5 hours), the assumption of $m = 2$ (linearly increasing) appears adequate for practical purposes, producing a very simple model. Earlier wearout would rise more rapidly and a nearly normal distribution ($m \approx 3.25$) is a very good approximation. Sometimes the infant mortality mode has the same shape (similar m's) for all mechanisms, resulting in a further simplification of the model as illustrated above. A more typical bathtub shape results when the curve is viewed on linear coordinates.

Undefined Mechanisms

An adequate model can be obtained when the mechanisms are unknown, assuming only two additive modes of failure. Suppose in the simulated data problem the mechanisms were not identified and parameters were estimated for the combined data (Fig. 2). Using the same procedure as for mechanism B, the following curve (in % per 1000 hours) was obtained:

$$h(t) = 0.024/t^{0.25} + 4.75t^{1.48}/10^9$$

This would suggest an infant mortality mode with shape parameter $m = .75$ and a wearout mode with shape parameter $m = 2.48$. Compared with the previous equation (Fig. 4), the curve is nearly identical beyond 1000 hours and differs only in the infant mortality region (dashed line - Fig. 4). The projected failure rates are very close but the estimated parameters are misleading because individual mechanisms were ignored. Thus, it is possible with this approach to fit an adequate "bathtub" model to a set of data by assuming two failure modes without a detailed knowledge of the failure mechanisms involved.

Summary

By viewing equipment or parts as a serial system in which any failure results in unit failure, the bathtub life curve is the sum of independent failure mechanism hazard rates. Evaluating each mechanism separately allows the parameters to be estimated readily by graphical methods. The complete model is then "built" additively from each mechanism model.

Very adequate bathtub curves result assuming zero location parameters which are easy to include, if necessary. In many practical situations, assuming a normal ($m = 3.25$) or a Rayleigh ($m = 2$) distribution is reasonable if rapid wearout occurs inside or outside the region of interest, respectively. These assumptions produce very simple models with which to monitor or project product performance lifetime.

It is not always practical to isolate and identify various failure mechanisms. However, it is possible to fit an adequate bathtub model by assuming two additive failure modes, realizing the resulting shape parameters are misleading.

References

1. K. H. K. Kao, "A Graphical Estimation of Mixed Weibull Parameters in Life Testing of Electron Tubes", Technometrics, pp. 389-407, November, 1959.
2. M. L. Shooman, Probabilistic Reliability: An Engineering Approach, New York: McGraw Hill 1968, pp. 193-194.
3. Wayne Nelson, "Hazard Plotting for Incomplete Failure Data", Journal of Quality Technology, Vol. 1, No. 1, pp. 27-52, January, 1969.

SOME INSIGHTS INTO STATISTICAL MODEL BUILDING*

Amrit L. Goel
Department of Industrial Engineering and
Operations Research; and Systems
and Information Science

Syracuse University
Syracuse, N.Y.

Summary

The primary purpose of this paper is to discuss several important aspects of statistical modelling which should be considered when developing such models. A brief description of the modelling process and the model development procedure is presented. An actual modelling example is used to illustrate the extra sum of square principle, the occasional need of transformation of the independent and the dependent variables, and to show the importance of properly interpreting the precision of the estimated parameters. The role of designed experiments in statistical modelling is also pointed out.

1. Introduction

The importance of mathematical models in various industrial economic and business investigations is well recognized. However, in many applications it is not feasible to obtain exact mathematical models and one needs to develop empirical models using statistical techniques.

The development of statistical models is not simply a matter of feeding the data into a computer and getting a beautifully formatted ANOVA table with a whole bunch of impressive looking statistics for a given set of hodge podge data. Anyone can get "a" model. But, statistical modelling is much more than that. It requires the active participation of both the statistician and the engineer in all phases of model development, including that of planning the data collection strategy. The latter is very important because no matter how sophisticated the analyses, there is not much that can be done with bad or inappropriate data.

The purpose of this note is to provide an insight into some of the aspects that should be considered when building statistical models. The intention is not to get into the theoretical and other details but simply to pinpoint the key elements and illustrate how they play an important role in determining which model is the right one.

The iterative nature of model building is first described in Section 2 followed by a brief overview of the model development procedure in Section 3. An actual modelling exercise is carried through in Section 4 to illustrate the important aspects alluded to earlier. The importance of correct interpretation of the precision of the estimated parameters is discussed in Section 5. Finally, a brief description is given in Section 6 of the role the designed experiments play in statistical modelling.

*This work was supported, in part, by RADC under Contract No. F 30602-71-C-0312.

2. Iterative Process of Statistical Modelling

Given a set of data which is subject to statistical variations, the basic aim is to find an empirical relationship between the independent and the dependent variables that satisfactorily describes the data at hand. There is no unique relationship that will satisfy this aim but the one that appears justifiable and reasonable to the investigator will be chosen. Obviously, the investigator does not know what this relationship is, but generally he has some feelings about it.

The development of the model is an iterative process. Before discussing this process, a distinction should be made between two possible situations that may occur: data have to be collected and data have been collected prior to modelling. In the former case an opportunity exists to plan an efficient experimental strategy so that the collected data can be easily analyzed. In the latter case, problems may arise if the data collection was not undertaken carefully.

The iterative process of statistical modelling is shown in Figure 1. The basic steps are: Postulation of a tentative model, estimation of the parameters and a check for the adequacy of the fitted model. Invariably, one has to go through the process of revising the postulated model several times before a satisfactory model is obtained. It is here that properly collected data and careful analyses help in reducing the number of iteration that will be required before getting the final model. If no satisfactory model is obtained, then the investigator may have to abandon the project or seek additional data.

3. Model Development Procedure

Investigators in various fields of applications are often concerned with the study of systems in which a dependent variable y is related to independent variables $x_1, x_2, \ldots, x_k$. As mentioned earlier, in many cases a theoretical relationship between the dependent and the independent variables is not known and one resorts to empirical modelling. Any such phenomenon can be represented as:

$$E(y) = \eta = f(\underline{x}, \underline{\beta}) \tag{1}$$

Where $\underline{x}$ are the independent variables and $\underline{\beta}$ are the parameters of the model.

In a great many applications, it is adequate to assume that the functional form f is linear in the parameters so that equation (1) can be written as:

$$E(y) = \eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k. \tag{2}$$

The observed values y are given by:

$$y_i = \eta_i + \varepsilon_i, \; i = 1,2,\ldots,n, \tag{3}$$

where ε_i are random with mean zero and some known probability distribution.

After postulating the model in equation (2), the investigator is faced with the following questions:

a) What are the estimates of the parameters $\underline{\beta}$?
b) How precise are these estimates?
c) Is the model adequate to describe the data?

There are several methods for obtaining the estimates of $\underline{\beta}$. The most common one is the method of least squares. This method yields the estimates $\underline{b}$ of $\underline{\beta}$ such that the sum of squares of the discrepancies between the observed values and the unknown mean is minimized. If we assume that ε_i in equation (3) are normally distributed with a constant variance σ^2 then $\underline{b}$ are also the maximum likelihood estimates of $\underline{\beta}$. The equation for getting $\underline{b}$ from the model in (2) is:

$$\underline{b} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{Y} \qquad (4)$$

Where $\underline{X}$ is the nxk matrix of independent variables, and $\underline{Y}$ is the nx1 vector of observations.

The precision of the estimates $\underline{b}$ is obtained from the equation

$$\hat{V}(b) = (X'X)^{-1}\hat{\sigma}^2, \qquad (5)$$

where $\hat{\sigma}^2$ is an estimate of the error variance σ^2.

As will be illustrated later, care must be exercised in interpreting the precision of the parameters.

To judge the adequacy of the model, a lack-of-fit test can be conducted if an independent estimate of σ^2 is available. Also, a careful study of the residuals from the fitted model is essential to determine any inadequacies in the model. Such a study also reveals whether additional terms should be added to the model or if any transformation of the variables is necessary to get a satisfactory model.

In addition to the above steps, the Analysis of Variance table is also carefully studied. Extra sum of squares contributed by each variable to the total regression sum of squares is evaluated to see how significant the addition of that variable to the model has been.

4. Numerical Example

The procedure outlined above and many of the points raised in the preceding paragraphs will be illustrated by considering an actual modelling problem. The problem deals with the development of a prediction model for the average fault location/checkout time, $\bar{T}_d$ as a function of $\bar{X}_T$ and H(s,f). $\bar{X}_T$ is a measure of test complexity and H(s,f) is a measure of information transfer from symptoms to failure. The data for the example have been taken from the ARINC[1,2] study and only the highlights of the analyses are discussed here. For detailed analyses the reader is referred to Goel and Barasia[5]. The data were collected by ARINC on 15 systems, 11 airborne and 4 ground, and are reproduced in Table 1.

4.1 First Order Model in $\bar{X}_T$ and H(s,f)

Let us postulate a first order linear model

$$E(\bar{T}_d) = \beta_0 + \beta_1\bar{X}_T + \beta_2 H(s,f) \qquad (6)$$

The estimates $\underline{b}$ for $\underline{\beta}$ are obtained from equation (4) where,

$$\underline{X} = \begin{bmatrix} 1 & 28.64 & 2.7302 \\ 1 & 158.6 & 2.91919 \\ 1 & 141.7 & 3.22059 \\ 1 & 109.9 & 2.8644 \\ 1 & 62.67 & 3.0497 \\ 1 & 137.6 & 2.7771 \\ 1 & 68.05 & 3.19736 \\ 1 & 205 & 5.08319 \\ 1 & 23.54 & 2.516 \\ 1 & 48.6 & 2.5191 \\ 1 & 16.95 & 5.636 \\ 1 & 15.7 & 7.5329 \\ 1 & 21.93 & 6.8270 \\ 1 & 59.76 & 4.7651 \\ 1 & 39.99 & 7.827 \end{bmatrix} \quad \text{and,} \quad \underline{Y} = \begin{bmatrix} 8.50 \\ 68.11 \\ 36.45 \\ 42.87 \\ 20.17 \\ 29.65 \\ 42.79 \\ 36.35 \\ 10.58 \\ 25.67 \\ 47.01 \\ 26.15 \\ 38.56 \\ 32.5 \\ 38.44 \end{bmatrix}$$

On substituting these values in equation (4), we get

$$\underline{b} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 11.51 \\ 0.14 \\ 2.77 \end{pmatrix}$$

and the fitted model is

$$\hat{T}_d = 11.51 + 0.14\,\bar{X}_T + 2.77\,H(s,f) \qquad (7)$$

The ANOVA Tables showing the contributions of $\bar{X}_T$ and H(s,f) to the regression sum of squares are given in Table 2. It is seen that neither of the independent variables is significant at 5% level ($F_{1,12;.05}=4.75$). However, $\bar{X}_T$ is significant at 10% level ($F_{1,12;.10}=3.18$) irrespective of the order in which $\bar{X}_T$ and H(s,f) are introduced in the model. Note that $\bar{X}_T$ and H(s,f) are not orthogonal and hence the contributions to the regression sum of squares depend on the order in which the two variables are introduced in the model. This distinction is very important in judging which variable is really significant from a statistical viewpoint.

4.2 Linear Model With Log-Transformation

A study of the residuals from the model in equation (7) indicates that transformation of the independent and dependent variables should be considered. Due to its simplicity and common usage, we first develop a model with log transformation of the dependent and the independent variables. The fitted model is

$$\ell n\,\bar{T}_d = 0.63 + 0.42\,\ell n\,\bar{X}_T + 0.80\,\ell n\,H(s,f) \quad (8)$$

The ANOVA tables for this model are given in Table 3. From the ANOVA Table in (a) we see that ℓn H(s,f) is significant at the 5% level ($F_{1,12;.05}=4.75$) if $\ell n\,\bar{X}_T$ has been included in the model prior to introducing ℓn H(s,f). Just the opposite result is seen in the ANOVA Table in (b) where $\ell n\,\bar{X}_T$ is significant if ℓn H(s,f) is already in the model. Furthermore, both $\ell n\,\bar{X}_T$ and ℓn H(s,f) are significant at 10% level in the first ANOVA, but this is not the case in the second ANOVA Table. This clearly demonstrates the need for considering the extra sum of squares contribution by each of the independent variables under investigation.

TABLE 1

DATA SUMMARY FOR PREDICTION MODELS

Observation #	System	$\bar{X}_T$	H(s,f) (bits)	$\bar{T}_d$ (min)
1	AIRBORNE	28.64	2.7302	8.50
2	AIRBORNE	158.6	2.91919	68.11
3	AIRBORNE	141.7	3.22059	36.45
4	AIRBORNE	109.9	2.8644	42.87
5	AIRBORNE	62.67	3.0497	20.17
6	AIRBORNE	137.6	2.7771	29.65
7	AIRBORNE	68.05	3.19736	42.79
8	AIRBORNE	205	5.08319	36.35
9	AIRBORNE	23.54	2.516	10.58
10	AIRBORNE	48.6	2.5191	25.67
11	GROUND	16.95	5.636	47.01
12	GROUND	15.7	7.5329	26.15
13	GROUND	21.93	6.8270	38.56
14	AIRBORNE	59.76	4.7651	32.5
15	GROUND	39.99	7.827	38.44

$\bar{T}_d$ = Fault Location/Checkout Time (average)
H(s,f) = Estimated average information in the joint occurrence of a symptom and a failure.
$\bar{X}_T$ = $\bar{X}_1 + \bar{X}_2 + \bar{X}_3 + \bar{X}_4$, where
$\bar{X}_1$ = Complexity of characteristics being measured
$\bar{X}_2$ = Rapidity with which the characteristics may be measured
$\bar{X}_3$ = Ease with which the characteristics may be interpreted
$\bar{X}_4$ = Availability of circuit points for test

TABLE 2

ANOVA TABLES FOR FIRST ORDER MODEL IN $\bar{X}_T$ AND H(s,f)

(a) Order of Introducing Variables: $\bar{X}_T$, H(s,f)

Source	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio
X_o	1.69216E4	1	1.692E4	94.39
$\bar{X}_T \mid X_o$	5.82990E2	1	5.830E2	3.25
$H(s,f) \mid X_o, \bar{X}_T$	3.46352E2	1	3.464E2	1.93
Subtotal	1.78510E4	3	5.950E3	33.19
Residual	2.15151E3	12	1.793E2	
Total	2.00022E4	15		

(b) Order of Introducing Variables: $H(s,f)$, $\bar{X}_T$

Source	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio
X_o	1.69216E4	1	1.692E4	94.39
$H(s,f)\|X_o$	9.16454E1	1	9.165E1	0.51
$X_T\|X_o$, $H(s,f)$	8.37697E2	1	8.377E2	4.67
Subtotal	1.78510E4	3	5.950E3	33.19
Residual	2.15121E3	12	1.793E2	
Total	2.00022E4	15		

TABLE 3

ANOVA FOR THE MODEL WITH LOG TRANSFORMATION

(a) Order of Introducing Variables: $\ell n\ \bar{X}_T$, $\ell n\ H(s,f)$

Source	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio
X_o	1.73126E2	1	1.731E2	92.63
$\ell n\ \bar{X}_T\|X_o$	6.79381E¯1	1	6.794E¯1	3.64
$\ell n\ H(s,f)\|X_o, \ell n\ \bar{X}_T$	1.28851E0	1	1.289E0	6.89
Subtotal	1.75094E2	3	5.836E1	312.3
Residual	2.24288E0	12	1.869E¯1	
Total	1.77337E2	15		

(b) Order of Introducing Variables: $\ell n\ H(s,f)$, $\ell n\ \bar{X}_T$

Source	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio
X_o	1.73126E2	1	1.731E2	92.63
$\ell n\ H(s,f)\|X_o$	4.78220E¯1	1	4.782E¯1	2.56
$\ell n\ X_T\|X_o, \ell n\ H(s,f)$	1.48967E0	1	1.490E0	7.97
Subtotal	1.75094E2	3	5.83E1	312.3
Residual	2.24288E0	12	1.869¯1	
Total	1.77337E2	15		

TABLE 4

ANOVA TABLES FOR LOG TRANSFORMATION MODEL. (AIRBORNE SYSTEMS)

(a) Order of Introducing Variables: $\ell n\ H(s,f)$, $\ell n\ \bar{X}_T$

Source	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio
X_o	1.21407E2	1	1.214E2	789.4
$\ell n\ H(s,f)\|X_o$	4.56455E1	1	4.565E¯1	2.968
$\ell n\ \bar{X}_T\|X_o, \ell nH(s,f)$	2.11044E0	1	2.110E0	13.72
Subtotal	1.23974E2	3	4.132E1	268.7
Residual	1.23037E0	8	1.538E¯1	
Total	1.25205E2	11		

(b) Order of Introducing Variables: $\ell n\ \bar{X}_T$, $\ell n\ H(s,f)$

Source	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio
X_o	1.21407E2	1	1.214E2	789.4
$\ell n\ \bar{X}_T\|X_o$	2.56609E0	1	2.566E0	16.68
$\ell n\ H(s,f)\|X_o, \ell n\bar{X}_T$	8.10894E¯4	1	8.109E¯4	0.0053
Subtotal	1.23974E2	3	4.132E1	268.7
Residual	1.23037E0	8	1.538E¯1	
Total	1.25205E0	11		

TABLE 5

ANOVA TABLE FOR THE MODEL IN EQUATION (10)

Source	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio
X_o	121.41	1	121.41	1271
$(1/\bar{X}_T)/X_o$	2.94	1	2.94	30.75
Subtotal	124.35	2	62.18	650.8
Residual	0.86	9	0.096	
Total	9.61515E4	11		

4.3 Linear Model with Log Transformation (Airborne Systems)

The models described to this point do not seem to be adequate for describing the data for fault location/checkout time. On studying the raw data and the residuals from the fitted models, it appeared that data for the ground electronics systems should be eliminated and efforts should be made to develop satisfactory models for the Airborne Systems alone. We first fit a model using log transformations. The fitted model is

$$\ell n \hat{\bar{T}}_d = 0.24 + 0.72 \ \ell n \ \bar{X}_T - 0.04 \ \ell n \ H(s,f) \quad (9)$$

The ANOVA Tables in Table 4 show that $\ell n \ \bar{X}_T$ is the only significant variable at 5% level. It is interesting to note that $\ell n \ H(s,f)$ is significant at 20% level, if it is the first variable to be introduced in the model. However, the extra sum of squares due to $\ell n \ H(s,f)$ given that $\ell n \ \bar{X}_T$ is in the model, is extremely insignificant. From these results it appears that $\ell n \ \bar{X}_T$ is the only significant variable that should be considered for regression purposes. Further evidence to this effect is provided by a study of the data plots (not included).

4.4 First Order MOdel With Power Transformation of $\bar{X}_T$ and $\bar{T}_d$ (Airborne Systems)

The above results strongly suggest that H(s,f) should be dropped from the model. Also, it appears that transformations other than the log transformations should be investigated. One such class of transformations is the power transformation of the independent and the dependent variables as proposed by Box and Cox[3]. A brief description of this method is also given in Goel and Barasia[5]. Therefore, a first order model with power transformation of $\bar{X}_T$ and $\bar{T}_d$ was postulated.

Using this method, the fitted model[5] is

$$\ell n \hat{\bar{T}}_d = 4.05 - 44.20/\bar{X}_T \quad (10)$$

which has an R^2 of 0.77.

From the ANOVA Table in Table 5, we see that the transformed $\bar{X}_T$ is highly significant. A plot of the residuals vs $\bar{T}_d$ is given in Figure 2. A study of these plots indicates that the model in Equation (10) is quite adequate.

5. Precision of the Estimated Parameters

It was pointed out in Section 3 that one of the questions an investigator wants to answer is how precise the estimates are. Information about the precision of the estimates is obtained by considering the expression in equation (5). For the model developed in equation (10), we have from equation (5):

$$\hat{V}(\underline{b}) = \begin{pmatrix} .0258 & -1.04 \\ -1.04 & 63.59 \end{pmatrix} \quad (11)$$

i.e. $\hat{V}(b_0) = 0.0258$, $\hat{V}(b_1) = 63.59$ and $\hat{Cov}(b_0, b_1) = -1.04$.

The individual $100(1-\alpha)\%$ confidence limits for β_i are given by $\{b_i \pm t_{\nu;\alpha/2} \sqrt{\hat{V}(b_i)}\}$. Therefore, the 95% confidence limits for β_0 and β_1 are (3.69, 4.41) and (-62.26, -26.18) respectively. These confidence limits should be interpreted with considerable caution. Due to the fact that β_0 and β_1 are correlated, it is incorrect to draw conclusions about the feasible values of β_0 and β_1 based on their individual confidence limits. Instead, the joint confidence region for the two parameters should be used for this purpose. The method for obtaining a joint $100(1-\alpha)\%$ confidence region is discussed in Goel[6]. Using this method, the 95% confidence region for β_0 and β_1 is plotted in Figure 3. The individual 95% confidence limits for the two parameters are also shown in this figure. The nature of the contour delineating the joint confidence region depends both on the sign and the magnitude of the correlation between β_0 and β_1.

A study of Figure 3 shows that points lying within the individual confidence limits are not always included within the joint confidence region and vice versa. For example, consider point A with coordinates $\beta_0 = 3.8$ and $\beta_1 = -50.0$. This point lies within the individual confidence limits for the two parameters but is not contained in the joint confidence region. Therefore, this set of values is not acceptable at the 95% confidence level. On the other hand, point B, which is not included in the individual confidence limits, is admissible because it lies within the joint confidence region.

6. Role of Designed Experiments

In the preceding discussion we saw that many complexities were involved in the analyses and interpretation of data because the independent variables were not orthogonal. This also led to a difficulty in interpreting the precision of the estimated parameters. One way to avoid these difficulties is to plan an experimental strategy prior to data collection so that the matrix $\underline{X}$ is orthogonal. This ensures that the contributions of the individual variables to the total regression sum of squares will not be affected by the presence or absence of other variables in the model. This, in turn, will make the interpretation of the ANOVA table much easier. Furthermore, the estimated parameters will be uncorrelated so that there will be no need to consider the joint confidence region. For more details the reader is referred to Goel[6].

7. Conclusion

We have presented an outline of some of the aspects that should be considered in developing statistical models. There are several other aspects which were not dealt with due to the limitation of space and scope. For example, residual analysis, which is an important part of model building, has not been discussed at any length. A Bayesian interpretation of fitted models was also not attempted. However, enough points have been brought out that will enable a prospective investigator to gain a good insight into the pitfalls of statistical modelling.

References

1. ARINC Research Corporation, Maintainability Prediction and Demonstration Techniques. RADC TR-69-356, Volume 1.

2. ARINC Research Corporation, Maintainability Prediction and Demonstration Techniques, RADC TR-70-89.

3. Box, G.E.P. and Cox, D.R., An Analysis of Transformations, Journal of the Royal Statistical Society, Series B, Vol. 26, No. 2, 1964, pp 211-252.

4. Draper, N.R. and Smith, H., Applied Regression Analysis, John Wiley, 1966.

5. Goel, A.L. and Barasia, R.K., Models for Maintainability Prediction, TR No. 72-5, Department of Industrial Engineering and Operations Research, Syracuse University, 1972.

6. Goel, A.L., Computer Based Statistical Modelling, Unpublished Notes, Department of Industrial Engineering and Operations Research, Syracuse University, 1972.

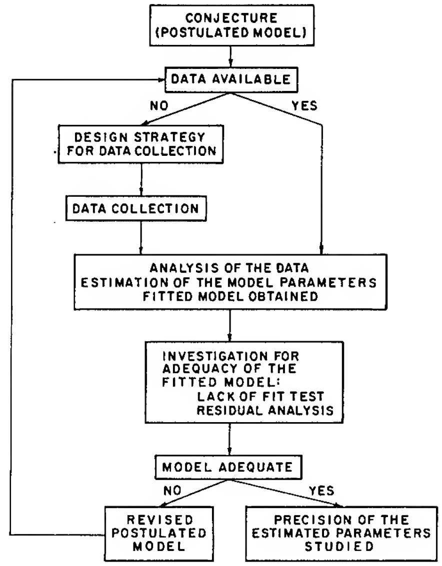


FIG. 1 ITERATIVE PROCESS OF STATISTICAL MODELLING

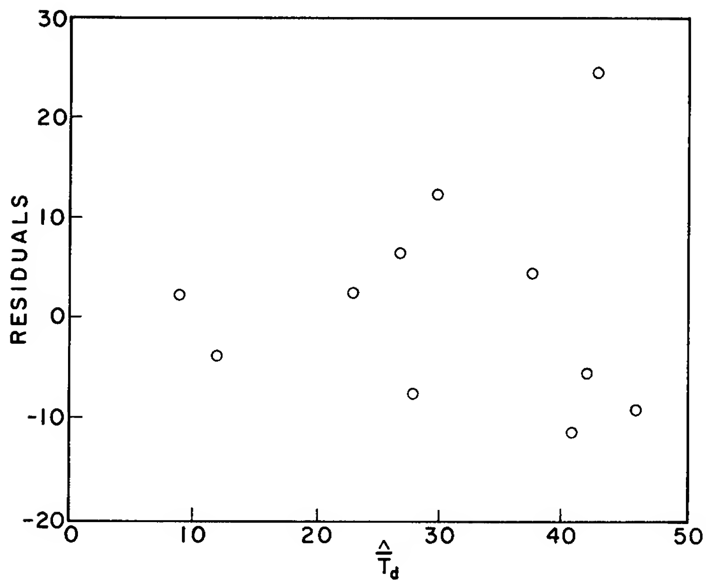


FIG. 2 PLOT OF RESIDUALS VS. $\hat{T}_d$

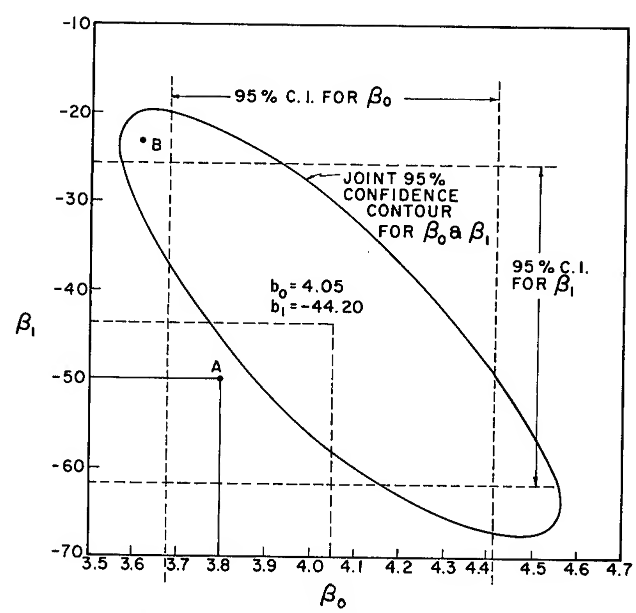


FIG. 3 CONFIDENCE REGION FOR β_0 AND β_1

MULTI BALANCE MEASUREMENT OF FLOW SYSTEMS

Shunsuke Nogita and Yoshihiro Uchiyama
Hitachi Research Laboratory of Hitachi, Ltd.
Omika, Kuji-machi, Hitachi
Ibaragi, Japan

SUMMARY

A fail-safe flow measurement technique is proposed. This "Multi Balance Measurement" takes advantage of material balances which generally exist in a steady state system. A screening criterion of balance inconsistency is applied to every balance group, and from the resulting combination of the groups having significant inconsistencies, an (or a few) erroneous measurement is detected and located in a set of simultaneously measured data. The measuring system with MBM works effectively with less redundancy.

INTRODUCTION

In process industries such as oil refinary, chemical, atomic power and so forth, long-period operation is a necessity for an economical plant. Though, for the purpose, various reliability techniques have been applied, there is still a strong demand for a new technique which will enable processes to be operated longer with low expenses. Besides, the introduction of computers to process control asks for new techniques which will take the place of experienced operators in many respects.

To meet the need a flow measurement technique, named "Multi Balance Measurement", will be presented in this article. This MBM is applied to a flow system, and locates a few erroneous measurements in a set of simultaneously measured data, providing for a fail-safe method with a smaller number of measuring devices.

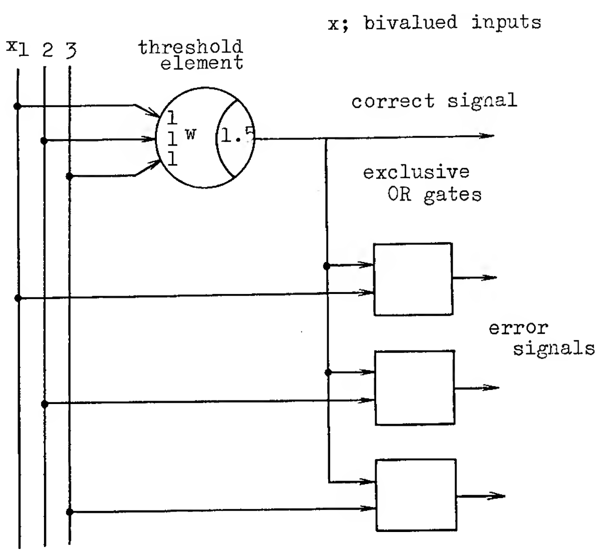


Figure 1. Example of Error Location System

As typically shown in Figure 1, an efficient realization of majority logic and error location to bivalued inputs is accomplished by a threshold element (1) and exclusive OR gates. Corresponding to this, Figure 2 indicates a realization of MBM, which intends to treat analog inputs (measured data) and locate a few measurements which contain systematic errors. The characteristics of MBM lies in (1) treatment of analog inputs, (2) the threshold derived

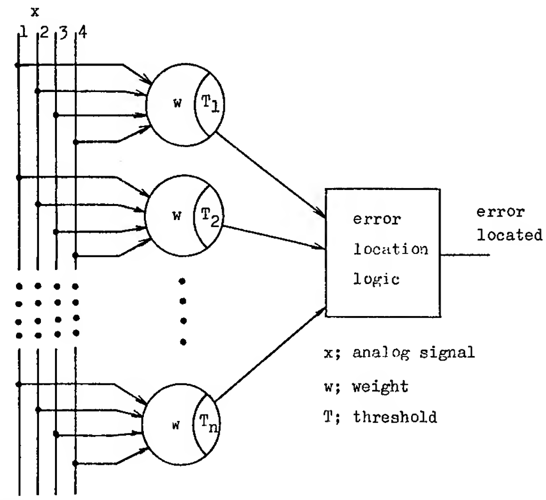


Figure 2. Signal Flow of Multi Balance Measurement

from statistical relations, and (3) the error location logic based on material balance equations.

OUTLINE OF MULTI BALANCE MEASUREMENT

Here a flow system is defined so that it has at least two components. Every component is connected to at least one other component by at least one route. Each component has at least one input route and one output route. From this definition at least three material balance groups can be composed in a system, if measuring devices are attached suitably to the routes.

And MBM is based on the following presumptions; (1) The system is in a steady state. That is, the following two relations, $E(x_i,t) = E(x_i)$ and $\mathrm{var.}(x_i,t) = \mathrm{var.}(x_i)$, exist in the system. (2) A (or a few) measurement in a set of simultaneously measured data might contain a (or a few) systematic error, while the others contain small random errors. (3) Every random error is statistically independent from every other. (4) The variances of random errors are previously known.

The process engineer would admit that in many cases these presumptions are acceptable for a real system.

Hence, the logic of error location is proposed as follows; When the deviation from a particular material balance is not significant, all the measurements in the balance group ($X_o \in \nu_o$) are assumed to be sound, i.e., contain no systematic error. Let us call the sound data set, whose elements are involved in all sound balance groups, ν_h.

$$\nu_h = \nu_{o1} \cup \nu_{o2} \cup \cdots \cup \nu_{op} \tag{1}$$

where suffix p indicates the number of sound balance groups.

In the same fashion, when the deviation from a particular material balance is significant, one (or a few) measurement in the balance group ($X_s \in \nu_s$) is assumed to be erroneous, i.e., contains a systematic error. Let us call the suspicious data set, whose elements are involved in all suspicious balance groups, ν_d.

$$\nu_d = \nu_{s,p+1} \cup \nu_{s,p+2} \cup \cdots \cup \nu_{s,n} \tag{2}$$

where suffix n indicates the number of material balance equations.

Then, the erroneous measurements are involved in the complimentary set of ν_h. That is, the erroneous data set ν_e is obtained from the following relations.

$$\nu_e = \overline{\nu_h} = \nu_d - \nu_h \tag{3}$$

SCREENING CRITERION

As was pointed out by Ripps (2), material balances are readily formulated in linear relations, if mass flow rates are adopted as variables. This linearity makes problems simple, and a clear-out criterion for screening inconsistency (threshold) was developed by one of the authors on statistical considerations (3).

The derivation of the criterion is briefly summarized in this chapter.

Suppose there are m measurements which contain neither random nor systematic errors, then it is possible to close a material balance. Let us call the ideal measurements, which will exactly close a material balance, t_i. The material balance is formulated by,

$$\sum_{i=1}^{m} a_i t_i = 0 \tag{4}$$

where a_i's are the balance coefficients of the system.

On the other hand real measurements generally contain either random or systematic errors. Let us call the real measurements, x_i. Then, a balance deviation can be defined by,

$$Z = \sum_{i=1}^{m} a_i x_i \tag{5}$$

The distribution of a balance deviation is the function of the distribution of each measurement. Assuming that the distribution of each measurement is normal, the distribution of a balance deviation is normal with a mean M_z and variance σ_z^2.

$$M_z = 0 \tag{6}$$

$$\sigma_z^2 = \sum_{i=1}^{m} a_i^2 \sigma_i^2 \tag{7}$$

where σ_i^2's are the variances of measurements.

The screening criterion of a balance deviation is derived from the above relations. The normalized deviation N, defined by

$$N = Z/\sigma_z \tag{8}$$

has a normal distribution with a mean zero and variance 1. Therefore, from a cumulative normal distribution table, for example, the probability of N being in the interval of -1.645 to 1.645 is read to be 0.90, that is, when $|N| > 1.645$ we might say that the inconsistency is significant with a type I error probability of 0.10.

EXAMPLES

The distinct understanding of MBM technique will be attained from the following examples.

Example 1

A simple system is shown in Figure 3. A and B are components, each of which has one input flow and one output flow. M1, M2, and M3 are flow meters attached to each flow.

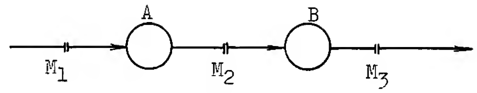


Figure 3. A Simple Flow System

In a steady state three material balance equations are acceptable.

$$t_1 - t_2 = 0, \quad t_2 - t_3 = 0, \quad t_3 - t_1 = 0 \tag{9}$$

where t_i's are the ideal measurements mentioned previously. The three balance deviations are defined by,

$$Z_1 = x_1 - x_2, \quad Z_2 = x_2 - x_3, \quad Z_3 = x_3 - x_1 \tag{10}$$

where x_i's are the corresponding real measurements.

For simplicity the variances of random errors are assumed to be the same and independent from one another. Let them be σ^2. Three normalized deviations result from Equation (10).

$$N_1 = Z_1/\sigma_{z1}, \quad N_2 = Z_2/\sigma_{z2}, \quad N_3 = Z_3/\sigma_{z3} \tag{11}$$
$$\sigma_{zi} = \sqrt{2}\,\sigma$$

where the variance of the normalized deviation is 1.

From Equation (11), the common measurement in two normalized deviation is easily found. For example, measurement x_1 is common in N_1 and N_3. Consequently, when there occurs the case of $N_1 > 1.645$, $|N_2| \le 1.645$, $N_3 < -1.645$, measurement x_1 is predicted to have a systematic error.

Table 1. Outcome and Judgement of Example 1

Outcome	Criterion			Probability	Judgement
	N_1	N_2	N_3		
1	0	0	0	0.8776	all sound
2	+1	0	-1	0.0046	x_1
3	-1	0	+1	0.0046	
4	+1	-1	0	0.0046	x_2
5	-1	+1	0	0.0046	
6	0	+1	-1	0.0046	x_3
7	0	-1	+1	0.0046	
8	+1	0	0	0.0158	not locatable
9	-1	0	0	0.0158	
10	0	+1	0	0.0158	
11	0	-1	0	0.0158	
12	0	0	+1	0.0158	
13	0	0	-1	0.0158	
14	+1	0	+1	0.0	
15	-1	0	-1	0.0	
16	+1	+1	0	0.0	
17	+1	+1	+1	0.0	
18	+1	+1	-1	0.0	
.	.	.	.	.	
.	.	.	.	.	
.	.	.	.	.	

+1; $N_i > 1.645$, 0; $|N_i| \le 1.645$, -1; $N_i < -1.645$

In this example there are twenty seven outcomes as partly shown in Table 1, depending on the state of each normalized deviation. Judgement of erroneous measurement is also shown. The probabilities listed in the Table, except for outcome 1, indicate the Type I error probabilities of rejecting the null hypothesis ($m_{z1} = 0$, $m_{z2} = 0$, $m_{z3} = 0$) when it is in fact true. These figures have been obtained from double integral of the probability density function over corresponding regions.

Example 2

A practical example of a chemical process is shown in Figure 4. Component A is a reactor. Two liquid flows, f_1 and f_2, enter into A, and one flow f_3 leaves A to a separator B. Vapor flow f_4 leaves from the top of B, and liquid flow f_5 leaves from the bottom.

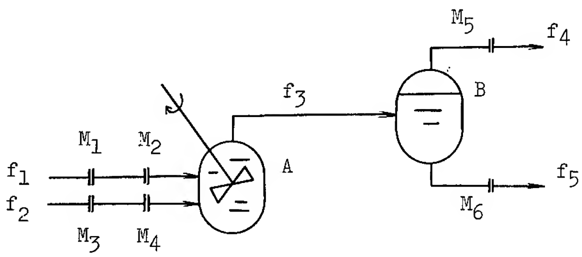


Figure 4. A Chemical Process

In this system the flow ratio of f_1 and f_2 should be kept constant to assure the stable quality of the reaction product. To make the measurement certain two flow meters are attached to each of the two input flows, f_1 and f_2. The flow rates of vapor and liquid from the separator are also measured, but the flow rate of f_3 is not measured because of the difficulty of measuring a two-phase flow.

For simplicity let the variances of random errors be the same.

$$\sigma_i = \sigma \quad (i=1,2,\cdots,6) \tag{12}$$

In a steady state the six material balance equations are

$$\begin{aligned} &t_1 - t_2 = 0, \quad t_3 - t_4 = 0 \\ &t_1 + t_3 - t_5 - t_6 = 0 \\ &t_1 + t_4 - t_5 - t_6 = 0 \\ &t_2 + t_3 - t_5 - t_6 = 0 \\ &t_2 + t_4 - t_5 - t_6 = 0 \end{aligned} \tag{13}$$

where t_i's are the ideal measurements.

The six balance deviations are defined by,

$$\begin{aligned} &Z_1 = x_1 - x_2, \quad Z_2 = x_3 - x_4 \\ &Z_3 = x_1 + x_3 - x_5 - x_6 \\ &Z_4 = x_1 + x_4 - x_5 - x_6 \\ &Z_5 = x_2 + x_3 - x_5 - x_6 \\ &Z_6 = x_2 + x_4 - x_5 - x_6 \end{aligned} \tag{14}$$

And the six normalized deviations which result from Equation (14) are

$$\begin{aligned} &N_1 = Z_1/\sigma_A, \quad N_2 = Z_2/\sigma_A \\ &N_3 = Z_3/\sigma_B, \quad N_4 = Z_4/\sigma_B \\ &N_5 = Z_5/\sigma_B, \quad N_6 = Z_6/\sigma_B \end{aligned} \tag{15}$$

where

$$\sigma_A = \sqrt{2}\,\sigma, \quad \sigma_B = 2\,\sigma \tag{16}$$

In this example there are $729(=3^6)$ possible outcomes in the combination of normalized deviation states. Among these, typical outcomes and corresponding judgements are listed in Table 2.

Table 2. Outcome and Judgement of Example 2

Outcome	Criterion						Judgement
	N_1	N_2	N_3	N_4	N_5	N_6	
1	0	0	0	0	0	0	all sound
2	+1	0	+1	+1	0	0	x_1
3	-1	0	-1	-1	0	0	
4	-1	0	0	0	+1	+1	x_2
5	+1	0	0	0	-1	-1	
6	0	+1	+1	0	+1	0	x_3
7	0	-1	-1	0	-1	0	
8	0	-1	0	+1	0	+1	x_4
9	0	+1	0	-1	0	-1	
10	0	0	+1	+1	+1	+1	x_5 or x_6
11	0	0	-1	-1	-1	-1	
12	+1	-1	0	0	0	0	not locatable
13	-1	0	-1	-1	-1	0	
.	.	.	.	.	.	.	
.	.	.	.	.	.	.	
.	.	.	.	.	.	.	

+1; $N_i > 1.645$, 0; $|N_i| \leq 1.645$, -1; $N_i < -1.645$

RELIABILITY OF MEASURING SYSTEMS

Let us take the same flow system as in Example 2 to show the effect of MBM on increased reliability of the measuring system. Here the mission of the measuring system is assigned to keep the constant flow ratio to the reactor.

(1) Without MBM

As there is no means to find an erroneous measurement when inconsistency between measurements x_1 and x_2, or between x_3 and x_4 exist, the reliability of the measuring system is

$$R_s = R_1 \cdot R_2 \cdot R_3 \cdot R_4 \tag{17}$$

because all measurements have to be sound.

(2) With MBM

When both measurement x_5 and x_6 are sound, the reliability of the measuring system, of at least one measurement being sound on each input flow, is

$$R_5 R_6 \{1-(1-R_1)(1-R_2)\}\{1-(1-R_3)(1-R_4)\}, \tag{18}$$

Even when at least one meter on flows f_5 and f_6 has failed, the system works if measurement x_1, x_2, x_3, and x_4 are all sound. The reliability of this case is

$$R_1 \cdot R_2 \cdot R_3 \cdot R_4 \cdot (1 - R_5 \cdot R_6) \tag{19}$$

After all, the reliability of the measuring system for keeping the constant flow ratio to the reactor becomes

$$\begin{aligned} R_{MBM} = &R_5 R_6 \{1-(1-R_1)(1-R_2)\}\{1-(1-R_3)(1-R_4)\} \\ &+ R_1 R_2 R_3 R_4 (1 - R_5 R_6) \end{aligned} \tag{20}$$

(3) Comparison

For simplicity assume the reliability of each component is the same R. From Equations (17) and (20), there results

$$R_S = R^4$$
$$R_{MBM} = (5-4R)\cdot R^4 \qquad (21)$$

The reliability of the measuring system with MBM is (5-4R) times larger than that of without MBM.

DISCUSSION

If statistical independency of random errors is eliminated from the presumptions, the following equation is introduced instead of Equation (7) (3). This gives a larger distribution of a balance deviation.

$$\sigma_z^2 = \sum_{i=1}^{m}\sum_{j=1}^{m} \rho_{ij} a_i a_j \sigma_i \sigma_j \qquad (22)$$

where the ρ_{ij}'s are the correlation coefficients between the i-th and j-th measurement.

This results in less sensitivity to the type II error probability, rejecting the hypothesis: the mean is different from zero, when that hypothesis is in fact true.

In Tables 1 and 2, indefinite judgements in some outcomes are not desired by the process analyst and the operator. Their situation could be improved in some degree if they use several sets of simultaneously measured data to increase the sensitivity of the screening criterion.

Although MBM stands on the above presumptions and detects an (or a few) erroneous measurement when it exists, if the operator is certain from other informations that the system is in a steady state and that the measuring devices are working correctly, inconsistencies in material balances can be interpreted as process disorder, which breaks material balances. This disorder detective capability is another advantage of MBM.

CONCLUSIONS

The technique proposed will find some application fields in process industries because of its practical presumptions, though previous knowledge of statistical parameters, variances and correlation coefficients of random errors, has to be provided from operator's experiances or at first from a rule of thumb. The MBM algorithm, when it is embodied in a computer control system, will be a good substitute for an experienced operator, which renders operational reliability to the process. The advantage of this technique in process economy will especially become clear when it is applied from the early stage of instrumentation planning.

ACHNOWLEDGEMENT

The authors wish to thank Dr. Shozo Shimada of Central Research Laboratory of Hitachi, Ltd. for helpful suggestions on error location problems.

NOTATION

a_i	material balance coefficient of a system
M	mean
m	number of measurement
N	normalized deviation
R_i	reliability of a component
t_i	ideal measurement which exactly close a material balance
x_i	real measurement which contains either random or systematic error
Z	balance deviation defined by Eq. (5)

Greek Symbols

σ_i^2	variance of random error
ρ_{ij}	correlation coefficient between the i-th and j-th measurement
ν_h	sound data set given by Eq. (1)
ν_d	suspicious data set given by Eq. (2)
ν_e	erroneous data set given by Eq. (3)

Mathematical Symbol

$\cup$	inclusive or

Subscripts

MBM	with MBM
S	without MBM
z	balance deviation

REFERENCES

1. Dertouzos, M.L., "Threshold Logic", The M.I.T. Press., Cambridge, Mass., 1965.
2. Ripps, D.L., Chem. Engng. Prog. Symp. Ser. 1965, No. 55, p8.
3. Nogita, S., Ind. Eng. Chem. Process Des. Develop., 11, No. 2, p197, 1972.

W. S. Oliveri
Electromagnetic Systems Division
Raytheon Company
6380 Hollister Avenue
Goleta, California 93017

H. J. Bailey, Jr. and A. Hollister
Conrac Division
Conrac Corporation
600 North Rimsdale
Covina, California 91722

INTRODUCTION

For purposes here, commercial technology is defined as an accumulation of theory, materials, process machinery, management, personnel expertise, production experience, and supporting software brought together in a commercial house to produce a reliable product repetitively. Theory differs from the other constituents in that it depends as much upon basic principles of physics in commercial as in military technologies.

Our present interest in commercial technology relates to its potential utility for the development of new components in new military applications, especially in the presence of currently limited military budgets and tight schedules. Development of new military components from the basic technology provides much greater flexibility than the more routine case of a military project selecting a mature commercial component for environmental qualification and adaptation. Development in this manner also has its problems. As will be shown, the concepts, problems and procedures, for applying a commercial technology differ greatly from those of qualifying a commercial component.

Two examples of components developed from commercial technology are discussed herein: airborne application of commercial hybrid microcircuits and ground application of commercial TV monitors. Each is of proven value in new military systems. The hybrid example reflects the viewpoint of the military contractor adapting the technology with support of the commercial producers. The TV example reflects the viewpoint of a commercial producer.

The outstanding conclusion of the present study is that adaptation of commercial technology for use in military applications is a complex task, in which it is difficult to control all of the variables necessary for reliable performance of hardware in military operational environments. Those attempting the adaptation may wish, before the task is complete, that they had employed a more traditional military developmental procedure. Nevertheless, the examples discussed show that the rewards, in terms of cost and scheduling advantages, can make the effort worthwhile.

Both the commercial and military assurance sciences have more than paid for themselves in making this technology utilization successful. Indeed we feel that without the discipline and methods of the assurance sciences that there would have been an endless stream of problems and pitfalls which could have caused the application efforts to be slowed down or even abandoned before success was attained. With the help of the assurance sciences, success included not only project completion, but also considerable savings in terms of time and money. One major assurance contribution was implementation of effective quality assurance, reliability, and maintainability programs based upon MIL-Q-9858, MIL-STD-785, and MIL-STD-470. A second major contribution was assistance to project offices in defining a suitable methodology for technology assessment, transfer, refinement, and updating in military applications.

EVALUATION OF COMMERCIAL UTILITY

Hybrid Application

The commercial technologies seem to have the greatest utility when the military application cannot be routinely satisfied by military components previously qualified. For example, when Raytheon's Electromagnetic Systems Division needed some compact, lightweight airborne digital circuits, hybrid technology was the suggested answer, but hybrids qualified to MIL-M-38510 were not available. A survey of hybrid sources, considering the above advantages, suggested that a good compromise (between qualification status and design factors) could be achieved. Major determinants were that multiple sources were available and failure rates were low enough to justify a throw-away maintenance concept. Side benefits included the availability of applicable computer aided design techniques for chip layout and thermal analysis; readily-available, hermetically-sealed, hybrid packages; and availability of multi-layer (up to 5 layers) ceramic substrates.

TV Application

The Conrac Division of Conrac Corporation is a strictly commercial producer of TV monitors for the network studios, various closed-circuit TV systems, and other commercial applications. The manufacturing and testing schemes for this operation have gradually been refined over the years to satisfy this specialized market with a quality product at competitive prices. The major military application for this product was in ground-based surveillance systems. The obvious initial advantage of the commercial TV technology was that it was already producing a reliable, functional product which could be adapted to a military system. Although the commercial operation could be organized to provide both commercial and military products, Conrac found it more feasible to provide the military monitors at its IC (Military) Division. This paper presents a comparison of the two operations,from the viewpoint of QA personnel in the commercial house.

Merits of Commercial Technology

One important reason for selecting a commercial technology as a candidate for a military application is availability, at relatively low cost, for inclusion in short developmental schedules. Other advantages, which became apparent during study of the hybrid and TV examples, included the following:

(1) Cost-Avoidance of research and development cost.
(2) Schedule-Hardware from the commercial technology is available off-the-shelf.
(3) Reliability-Failure rates and modes are established.
(4) Maintainability-Maintenance manuals and repair times are established.

(5) Quality Control - Inspection procedures are developed and debugged.
(6) Testing - Test capability, techniques, and instrumentation are available.
(7) Training - Experienced factory personnel are available for consultation.
(8) Safety - Hazards have been identified and eliminated.

The present study of these advantages is considered important not mainly for the particular examples discussed, but rather for the experience in efficiently using competitive commercial technology. Commercial as well as military houses have had to make their technology pay in order to stay in business.

Disadvantages of Commercial Technology

One disadvantage of commercial technology is the problem of controlling a large number of design and manufacturing variables which have not been defined by a military specification. In the cases cited, the initial concern was that the hybrid IC and TV monitor products were not rugged enough to withstand the environments specified for the military applications. Such concern later seemed minor compared to:

(1) Lack of product design, manufacturing, and test documentation in accordance with military criteria.
(2) Lack of standardization for interfacing with associated military systems.
(3) Maintenance problems, because military personnel were not familiar with this kind of equipment.
(4) Logistics problems, caused by extensive use of non-standard parts and components in the commercially adapted equipment.

As discussed in the examples, these disadvantages were not fully overcome. Rather, new methods of documentation, design, and maintenance were introduced. The assurance sciences, of course, played a key role in making the indicated compromise successful.

ROLE OF ASSURANCE SCIENCES

Four-Step Process

The four-step process of adaptation as defined here (technology assessment, transfer, refinement, and updating) is directly applicable to the hybrid and TV examples studied and may be adaptable to many other applications. Table 1 shows some of the reliability factors associated with these steps in the hybrid example.

Technology assessment in this context involves literature searches, facilities surveys, and any other evaluation techniques needed to define critical characteristics of the technology. It provides the basic data for any necessary design changes and for reliability program support. Transfer is the process of selecting specific parts of the technology (e.g. hybrid packaging) to be used in the military application. The refinement phase of the adaptation process is defined as similar to the usual military design and development program, using reliability tools such as failure mode and effect analysis, evaluation testing and other applicable portions of MIL-STD-785. Updating refers to the usual problem of replacing materials, processes, and parts of the technology which gradually become obsolete.

Reliability Program

Although the role of the assurance sciences in the hybrid IC and TV monitor examples was found to involve greater-than-usual support, it was deemed advantageous to use standard reliability engineering techniques as a starting point. The practical program for both examples reflects the major tasks of MIL-STD-785 adjusted for adaptation phases as shown in Table 2.

Assessment

Personnel carry-over (i.e., the present employment in military systems houses of personnel previously experienced in the applicable commercial technology) is characteristic of both the hybrid and TV examples. The presence of these personnel has simplified the task of technology assessment for their applications. Their experience has provided insight into potential problems and reliability factors at all phases of the technology adaptation.

Literature searches and facilities surveys have provided the initial data for reliability characterization of each element to be utilized. (GIDEP and FARADA have also been most useful for determining failure rates and failure modes.) The initial data is important for defining the adaptation program, and for identifying the alternative courses of action to be evaluated. In the hybrid example, this included evaluation of chip masses to minimize susceptibility to aircraft vibration, and evaluation of circuit layouts to minimize pinouts and cross-talk. In the TV monitor example, this included mainly evaluation of packaging to minimize effects of transportation and humidity environments. In both examples, the early assessments indicated a degree of incompatibility between existing interconnection methods and standard mil-spec items. Forcing the technology to change, in order to accept the standard mil-spec connectors, would have offset much of the cost and scheduling advantage, so initial planning included evaluation testing of non-standard connectors.

Transfer

As might be guessed, with the actual transfer of these technologies came an awareness of the many loose ends which had to be resolved. The systematic approaches of the assurance sciences were very useful for keeping the work coordinated. Effective partitioning and packaging of the hybrids illustrate a few of the reliability variables.

(1) Partitioning. The initial circuit design partitioning embodied building blocks which were functional entities, in order to facilitate functional testing. However, this resulted in an unequal number of pinouts per package. In two cases, the number of pinouts exceeded the selected case limitation of 30 pins. The problem was resolved by moving suitable portions of the circuitry into other packages and adjusting the test sequences accordingly.

(2) Evaluation Testing. Because the hybrid IC's were digital functions, the usual problem existed of exercising all states during evaluation testing. At each state, the variable

TABLE 1

RELIABILITY FACTORS IN THE HYBRID EXAMPLE

(MILITARY CONTRACTOR VIEWPOINT)

	ASSESSMENT	TRANSFER	REFINEMENT	UPDATING
Materials	Standard packages, chips resistors, and leads available	"General" specification defines practical constraints upon choice of material	Chip size kept small to minimize vibration effects	Material updating will be controlled
Process Machinery	Functional test equipment and laser resistor trimming devices available	Specification requirements are compatible with available test and processing capability	Added high frequency clock capability to test equipment to test one hybrid device type	No problems anticipated
Management	Commerical management experienced via years of producing for commercial computer industry	Management advantage is retained by hybrid production at the commercial house	Additional control/review techniques added as necessary to meet military requirements	Standard program review techniques, per MIL-STD-785, will guide management of updating tasks
Software	Many drawings available although not to mil-spec format requirements. Computer aided design software available	Basic set of mil-spec drawings required	Some tolerances tighter than previous. Defect classification added	No problems anticipated. Has become typical mil-spec control
Personnel Expertise	Facility surveys demonstrate presence of artisans for layout and packaging	No action required. Product is produced and tested at manufacturer's facility	Some retraining required to meet MIL-Q-9858	Routine enforcement of MIL-Q-9858 requirements and changes thereto
Production Experience	Commercially oriented hybrid production lines observed in the facilities surveys	No action required. Product is produced and tested at manufacturer's facility	Refinement includes augmented testing	Manufacturer must be responsive to customer changes

TABLE 2

RELIABILITY PROGRAM VERSUS ADAPTATION TASKS

MIL-STD-785 Standard Tasks	Adaptation Tasks
5.1 Reliability Management Subcontractor Control Program Review	Assessment Personnel Carry-over Literature Search Facilities Survey Maintainability and Human Factors Tradeoffs
5.2 Reliability Design and Evaluation Design Techniques Predictions Parts Reliability Failure Mode Analysis Critical Item Control Effects of Environments Design Review	Transfer Design Definition Evaluation Specification Development Worst Case/Thermal Analysis Refinement Drawing Development Design Review Failure Analysis/Corrective Action
5.3 Reliability Testing and Demonstration	
5.4 Failure Data	
5.5 Production Reliability	Updating Quality Control Change Control

number of digital loads and leakage paths tended to create a different value of input loading on the test equipment. Fortunately, reliability personnel were able to utilize these data for reliability evaluation ot interior portions of the circuitry. It was found, for example, that the acceptance limits on the input load for a given state could be set to detect abnormal leakages of interior elements. Semi-automatic test equipment (already developed by suppliers) was utilized for these tests, which include evaluation of more than 200 separate states on the more complex hybrid IC's.

(3) Specification Development. The major reliability influence was upon the general specification, wherein it was necessary to control package materials, leads, and other design and construction variables. In this instance, the practice of controlling reliability by prohibiting the supplier from incorporating any change in design, processing, or materials had to be implemented with care. As might be expected, commercially-oriented suppliers do not readily accept such restrictions.

(4) Mechanical Design. The design configuration for mounting the hybrid IC's to the PC Boards provided low thermal resistance between the hybrid IC and the PC board, since the equipment was conduction cooled. Another constraint on the mechanical design was that the mounting configuration had to facilitate flow soldering of the PC boards once the hybrid IC's were mounted.

Updating

Updating of hardware developed from commercial technology appears no more difficult than that presented by normal obsolescence of hardware in military systems. In the hybrid example, semiconductor chip manufacturing procedures may change as much as semiconductors have in the past, but suitable replacements can be defined for logistics purposes. Likewise, for the TV technology, updated specifications and changes in procedures can be developed for unique TV hardware such as CRT's, yokes, and flyback transformers.

Refinement

For the hybrids, reliability participation in design reviews was a major task of continuing value for refining the technology in the military application. In a sample review, reliability inputs were required for evaluation of beam lead versus chip/wire construction, definition of the development/test/delivery sequences, evaluation of low power logic, evaluation of silicon nitride passivation to minimize contamination, and evaluation of existing test stations on a production line.

For the TV example, refinement of the technology required QA/reliability participation for two purposes: establishment of the basic technology at a different facility and different operating division of the company, and establishment of a manufacturing/procurement system in accordance with military requirements. The following comparison of QA practices at the two divisions illustrates the extent of participation.

Quality Engineering. This comparison of the two Conrac Divisions emphasizes that the "strength" of quality control at the commercial facility is not necessarily less than that at the military facility. However, the military operation does embody more formality and documentation. It was the prime responsibility of Quality Engineering to interpret the customer requirements with respect to formality and documentation, and to assure inclusion of these requirements into specifications, formal procedures, and operating practices. One example was control of scratches on CRT faceplates. Because the military requirement was more stringent than the commercial requirement, it was necessary to reflect the requirement in procurement specifications, receiving inspection criteria, and handling procedures. Another example was certification. The commercial division commonly certifies the end-item product against functional and test specifications. The military division also found it necessary to certify compliance with receiving inspection tests and with raw material specifications. A third example was a change-over from commercial electronic parts to military specification parts. However, because a good grade of commercial parts was already being used, mil-spec versions of these parts were readily found. (Some redesign,of course, was unavoidable). A fourth example was implementation of the formality of MIL-Q-9858.

Reliability Engineering. In the military facility, reliability engineering tasks per MIL-STD-785 differ greatly from commercial efforts, although aims of the program differ little. For example, detailed math modeling and predictions per MIL-HDBK-217A have the same intent as commercial predictions of overall product reliability, which are based upon in-house data, customer data, and warranty costs. The bases of these predictions, in both cases, include part failure rates and knowledge of system operating environments. The outputs are identification of high-failure-rate items and data for reliability/cost tradeoff studies (which may receive much greater attention in commercial products than in military products). Nevertheless, the great amount of product analysis, per MIL-STD-785 requirements, probably does eliminate some product faults and thereby produces higher reliability in the field. Examples include detection of unnecessarily critical failure modes (by formal failure mode analysis) and detection of circuit hot spots or instabilities (by worst-case circuit analysis). Other formalities of the military reliability program include use of reliability program plans, and deliverable data. One reliability task which is not noticeably different is the collection and use of failure data. At the commercial division, in-house and field failure rates of parts, modules, and PC boards were closely monitored in order to minimize costs for TV test, rework, and warranty repair.

In-Process Manufacturing Control. In-process control at the two facilities is similar, except for the greater formality and documentation in military manufacturing. In both cases, work stations are laid out, equipped, and staffed to provide a controlled flow of hardware, with QC inspections at all critical stages. One key difference in formality is the military provision for in-process inspection by the customer (before completion of manufacturing steps which would "cover-up" potential defects). In the commercial facility, it was found that such customer inspections are unnecessary because of an existing strong desire to eliminate hidden defects. Such defects cost time, money, and customer goodwill in commercial facilities.

Test and Alignment Control. The military adaptation of TV circuits presented some interesting problems in test and alignment control, due to interactions of signals. In TV, there are relatively critical relationships for phasing of sync signals, control of feedback signals, regulation of high voltages, and impedance matching. Commercially, we have established routine procedures for set alignment and test, but recognize the need for "artistic refinement" of the procedure when routine alignment fails to produce the desired picture quality. The refinement is considered typical of the fine tuning done for many pieces of RF equipment, whether commercial or military. Other similar "problems" were control of high voltage arcs and transients and control of inductive component dimensions and lead lengths. In this case, the transfer of key personnel from the commercial to the military division provided the necessary expertise for minimizing the problems. The availability of these personnel, of course, greatly minimized demands upon military schedules and budgets. Experience indicates that refinement of totally new RF systems otherwise can be time consuming and costly.

Control of Measuring Equipment. Here, too, the main difference between the commercial and military facilities is the degree of formality and documentation with minor differences in accuracy. What is essentially a MIL-I-45208 inspection system is available commercially for measuring equipment and other quality tasks. In the military facility, this becomes a MIL-C-45662 Calibration System for concurrent use with a MIL-Q-9858 quality system. The more formal system requires (a) direct control of all test equipment utilized for acceptance testing, (b) use of standards traceable to the National Bureau of Standards, (c) use of history records on each piece of test equipment, and (d) use of measurements ten times more accurate than that required for each parameter measured. Commercially, work is done to whatever standard gets the job finished at a competitive cost. For many measurements, the standard is less stringent than the military; for others, it is more stringent. However, since the more formal system was already in effect at the military division, its implementation on the TV production lines was routine.

Supplier Surveillance. The main difference between military and commercial supplier surveil-

lance may be compared to input versus output contracting. On the one hand, it is assumed that imposition of quality controls, quality tests, and quality audits (inputs) on the supplier facility will produce the desired product quality. On the other hand, it is assumed that a good historical record plus competitive reasons (outputs) are assurance that the supplier will deliver the desired quality product on time. The change-over from one system to the other obviously did introduce problems for procurement of TV speciality items, but these were gradually resolved by compromise (or selection of new suppliers).

CHARACTERISTICS OF COMMERCIAL TECHNOLOGY

As indicated above, one of the major findings in the adaptation tasks, for both the hybrid and TV examples, was the gradual realization that the problems were complex. Initially, both cases were attacked as if a simple environmental qualification of hardware was the major reliability task. In retrospect, we find that the following major variables require consideration.

Materials

Our studies show that commercial materials are generally popular materials which competition has made available in large quantities at low cost. They usually can be easily machined, even automatically,although they are procured with somewhat less stringent requirements on dimensional tolerances and physical properties than are usually imposed on military materials. In some cases, however, the less stringent requirements may be adequate because they represent a practical empirical design solution versus the worst-case design approach typically selected for military applications. Another advantage of the commercial materials is that they are supported by a large body of historical performance data including publication of demonstrated characteristics in trade literature and manuals (ASTM, etc.). In addition, practical machinery and procedures are already developed for materials processing, while undesirables (e.g. materials presenting safety hazards) have been minimized. Such materials in the case histories discussed herein include high-voltage wiring for TV monitors, and Kovar case materials for hybrid IC's.

Process Equipment

The main commercial process equipment of interest to military application of technology include the semi- or fully automatic assembly and test equipment (although jigs and fixtures in a given technology are also economically important). This equipment has been debugged in commercial usage and can be easily maintained, thereby minimizing maintenance downtime and the danger of slipping production schedules. It is readily available and relatively inexpensive, especially at present, since the electronics industry is not operating at full capacity. Commercial process equipment for the case histories discussed include beam lead bonders for hybrid IC's and special fixtures for testing of high voltage TV monitor components.

Managers

Of necessity, managers of commercial technology have already made it as efficient as possible, thus debugging it for potential military applications. These managers are adept at using their resources effectively because they:

(1) Have withstood the test of both domestic and foreign competition.
(2) Have satisfied their stockholders with respect to control of cost factors which impact profits.
(3) Are generally results-oriented.
(4) Are generally cautious about taking any risks by changing their technology.

Without adequate and direct management experience, several military systems firms have lost money tackling hybrid technology.

Software

Another attractive feature of commercial technology in this sense is the availability of software (valid documentation and debugged computer programs) to support its ready adaptation to military application. Documentation available includes drawings, specifications, schematics, block diagrams, wiring lists, test procedures, and field support manuals. (To meet full mil-spec requirements, it may be necessary to reformat this documentation.) Examples of available computer programs include those for generating wire lists, semiconductor test programs, and numerical control machine instructions. These are important because, in many digital systems, the resources required to develop software can exceed the resources required to develop the hardware. Commercial software often requires little or no modification for the military application because the application differences are usually minor. For example, it does not have to be modified for environmental differences.

Personnel Expertise

Two types of personnel expertise in commercial technology are considered. One is the normal years-of-experience with applicable tools and fixtures. The other is the contribution of artisans. It is the artisans who somehow make the material, machine, or product functional when normal specifications and written procedures do not get the job done. Their on-hand data for responding to practical questions concerning cost and schedules may be limited, but they can intuitively define effective corrective action when failures occur. For either type, personnel expertise includes a familiarity with properties of commercial materials, as well as working knowledge of set-up and maintenance requirements for the associated process machinery.

Production Expertise

The existence of a smooth-running production line embodying an applicable technology is considered to be as valid a testimony of the utility of a technology as the routine case of qualifying a commercial component for a military application. The number of changes and additions for military adaptation of a product are greater and more complex than qualifying a component. One hazard in extensive change is that the stability of the original commercial production process may be severely upset before successful adaptation for the military application is achieved.

Hybrid Example

In this military jet aircraft application, it was necessary to quickly increase the capability of two digital circuits which were packaged as 14 printed-circuit boards using monolithic IC's (SSI), while simul-

taneously reducing the physical size of the circuits so that they could be packaged as 3 printed-circuit boards. Hybridization appeared to be the most logical solution to this design problem.

The customer was concerned about the addition of more non-standard parts to the system as well as the reliability risk associated with newly developed parts. He had seen very little historical data with which to assess the magnitude of this risk. The customer had, on previous programs, experienced schedule slippages and cost overruns due to the failure of some newly developed components to perform in accordance with specified requirements.

To preclude the occurrence of this type of problem, a systematic survey was conducted to identify manufacturers of similar hybrid devices from among dozens of possible candidates. Proposals were then requested from about a dozen manufacturers of similar hybrid devices. A source selection board with proper representation from Engineering, Manufacturing, and the Program Office was set up to evaluate these proposals. The major criteria for selection were technical approach, management, manufacturing, quality assurance, cost, and product support.

The proposals were based on specifications which firmly defined the environmental, functional, mechanical, and other requirements for the devices. All of the general requirements as well as the quality and reliability assurance requirements were established by a general specification. Detail requirements, specific characteristics, and other provisions unique to a particular device were specified in eight detailed specifications and their associated electrical schematic diagrams.

The biggest problem encountered in this example was testing. The complexity of these devices made comprehensive manual testing impractical. Automated testing required programming of the automatic hybrid device tester. The problem was how to debug the program which would be used by the device tester. The solution to this problem was to build breadboard circuits out of discrete components and use these breadboards to debug the program.

Another problem was assuring the customer that the devices were reliable enough for this application. For this reason, the following requirements were incorporated into the specifications:

(1) A maximum allowable failure rate which was to be demonstrated by analysis using best available data.
(2) The case to semiconductor junction temperature rise which was limited to 25°C in order to eliminate local hot spots.
(3) Process conditioning, testing, and screening of 100% of the devices. This included high temperature storage, mechanical shock, pre-cap visual inspection, seal leak tests, temperature cycling, burn-in, and electrical tests.
(4) Solderability and life tests on a sample of the devices.

So far the program has been successful, i.e., all problems have been resolved.

TV Example

Many variables in the TV example have been listed above in the discussion of technology refinement. In summary, the major steps for use of this commercial technology in military applications included the following:

(1) A decision to produce the hardware at the Instrument/Control (Military) Division rather than at the Conrac (Commercial) Division, because it was more cost effective to transfer a few key personnel, equipment, and procedures than it would be to change over the existing commercial production procedures and documentation systems.
(2) Specification development, reflecting customer requirements at all levels of hardware procurement and build-up.
(3) Supplier development, especially for suppliers of TV specialty items.
(4) Electronic parts change-over, from commercial to military, and subsequent circuit redesign where required.
(5) Repackaging, to withstand military environments (mainly shipping).
(6) Development of alignment and test routines, based upon commercial expertise.
(7) Implementation of normal military requirements (mainly the assurance sciences) for product manufacturing, test, inspection, and identification.

SUMMARY

There is a place for commercial technology in military applications. The concepts for applying a commercial technology differ greatly from the concept of qualifying a commercial component. The assurance sciences play a key role in making this technology utilization successful. In military applications this role involves defining a suitable methodology for commercial technology assessment, transfer, refinement, and updating.

Military airborne application of commercial hybrid microcircuits and military ground application of commercial TV monitors have been discussed as proven examples of commercial technology utilization. The hybrid microcircuit example illustrates the case of the military contractor adapting the technology with support of the commercial manufacturer. The TV monitor example demonstrates the case of the commercial manufacturer adapting the technology directly to a military application. In both cases, the task was complex, but the cost and schedule rewards made the effort worthwhile.

AVAILABILITY DEMONSTRATION RISK ASSESSMENT
– and Its Relationship to the Queue M/M/1*

T. W. Hasslinger
Radiation, A Division of Harris-Intertype Corporation
Melbourne, Florida

Acknowledgements

The writer would like to begin by noting the numerical analysis contribution of Mr. Klaus Mahn in providing solutions to the inverse risk function developed herein.

Summary

Producer's risk is treated in the demonstration of availability for large systems. A fixed time demonstration technique is discussed and an accept criterion identified. A risk function is developed and tested for its asymptotic behavior. From this function, an availability risk graph is presented which is suitable for graphical solutions. Examples of graphical solution techniques are provided. Finally, relationships between availability and queueing theory are examined in the appendix.

Introduction

Over the past several years system owners have been placing increasing emphasis on the actual amount of time equipment is capable of performing its task. This emphasis is due to increasing consolidation of tasks to a single equipment, increased criticality of information, escalating information rates, and growing cost of ownership. It is axiomatic that such emphasis would manifest itself in the form of proof upon purchase that an equipment will perform to minimum availability requirements. This paper addresses the problem of producer's risk in such a demonstration.

The paper begins by defining a demonstration technique, after which a risk function is developed and a new formulation of availability is offered. System down time, rather than availability itself, is selected as the decision variable. The theme is that of providing a ready means of a priori demonstration risk assessment by treating a system under test as a single, infinite source of failures which can queue up to be repaired. This notion is believed to afford a much more viable model of large system availability and lends itself to analysis by queueing theory. The model developed here assumes a single maintenance facility with exponential service and interarrival time distributions. The appendix shows that availability problems under these conditions can be treated using theory for the Queue M/M/1.**

For the benefit of the reader who may not feel comfortable with queueing mathematics, none of the development or derivation in the paper relies on queueing theory. The paper can in fact be read with no knowledge of queues as all references to queues are relegated to the appendix. Only two points in the text are made with expressions drawn from the appendix without proof.

Demonstration Description

While the purpose of this paper is to develop an availability risk function and discuss its implications, it should be obvious that an arbitrary risk function can have little practical appeal. A risk function can only take on significance in context with a designed test. This section will summarize the characteristics of the developed availability demonstration. Subsequent sections will treat each characteristic in detail and provide further qualifications.

Test Duration

The test is designed to operate for a fixed time (see **Development of the Risk Function** for further qualifications). Test time is arbitrary so long as it is greater than approximately ten times the system mean restoration time.

Accept Criterion

The system under test will be accepted as meeting specification if the accumulated down time is equal to or less than the product of specified unavailability and test time. The system is otherwise rejected.

Down time has been selected as the decision variable since it is directly measurable and leads to a more efficient test. If availability was used as the decision variable, two means of determining this quantity may be used: (a) calculate availability from system up and down time, (b) sample the state of the system over time and develop a binomial distribution. At this writing, neither approach seems appealing but the latter would make an interesting paper.

Applicability

The test is designed to demonstrate availability for large systems with a continuous demand for use. Message routing systems, computer complexes, and communication satellite terminals are examples of such systems. Smaller systems can, of course, be demonstrated with this technique, but conventional approaches may well prove satisfactory.

System State

Since the system under test is treated as a single, large source of failures, any failure is assumed to place the system in a failed state and the system is allowed to enter lower states within the failed state (see **State Description**). In this regard, accountable failures must be carefully defined. An element failure within a redundant network will likely not be chargeable as a system failure. It is not uncommon, however, to treat a redundant system as a hypothetical single string for purposes of demonstration. This eliminates much of the confusion over accountable failures and typically shortens the demonstration. When the single string approach is taken, the availability requirement must be adjusted downward to reflect this artificial, albeit practical, situation.

Service Policy

The test is designed to treat repair of failures in a sequential manner, with no more than one repair action at any one time. And, from the preceding discussion, the system is allowed to fail while being repaired. The system will thus be down until each of the failures is worked off one by one. This policy is not as radical as it may first appear. Systems are often maintained with a single repair crew or are provided with test equipment or diagnostics which will allow only a single repair action at a time. In addition, a second failure in a large system may go undetected until the first failure is repaired.

*The Availability Risk Graph appearing as Figure 3 is published with permission of Harris-Intertype Corporation.

**The notation M/M/1 is a queueing classification due to Kendall: M – exponential interarrival times/M – exponential service time/1 – single server.

Equilibrium Availability

Availability is defined over all time, beginning with initial conditions forced upon the system and, if the system exhibits a stationary distribution, transitioning to an equilibrium which is time invariant. This demonstration technique is designed to test specified equilibrium or steady-state availability. The influence of the transitional behavior is, however, included in the risk function.

Assumptions

Failure Distribution

The time between failures is assumed to be exponentially distributed. This assumption should only be bothersome in the case of redundancy. For the class of systems addressed here, the chances are great that such redundancy will be maintained. McGregor[1] shows that, under a broad range of repair and failure rates, maintained redundancy with immediate repair behaves exponentially. Einhorn[2] shows similar results using a different approach.

Restoration Time Distribution

It is assumed that the time to restore is exponentially distributed. While it is well recognized that restoration time is lognormally distributed, a reasonable set of lognormal distributions can be classed as "exponential enough" for practical purposes. Goldman and Slattery[3] have collected restoration time data to indicate that the standard deviation (σ) of the transformed normal (from the lognormal) ranges between 0.6 and 1.4 for most electronic systems. These data are admittedly old and the effects of MSI and LSI have yet to be determined. However, large systems still contain a generous mixture of electronic and electromechanical devices. Furthermore, MSI, rather than reducing the size of systems, is allowing systems with more functions and sophistication to be built on the same floor space.

If a system is completely modular with automatic fault isolation, σ should be close to 0.6. For large electromechanical systems or systems which require manual isolation to the piece part, σ should be close to 1.4. For the class of systems considered here, a value of unity for σ is reasonable. Figure 1 shows cumulative distribution plots of the exponential and the lognormal for three values of σ. The plots are normalized on the means of the distributions; i.e., they all have the same mean. Since the paper is log-probability, the lognormal plots as a straight line. Note from the figure that there is little practical difference between the exponential and the lognormal with a standard deviation of unity. Note further that the exponential distribution crosses each of the lognormal plots at two points, thus providing some degree of bilateral correction.

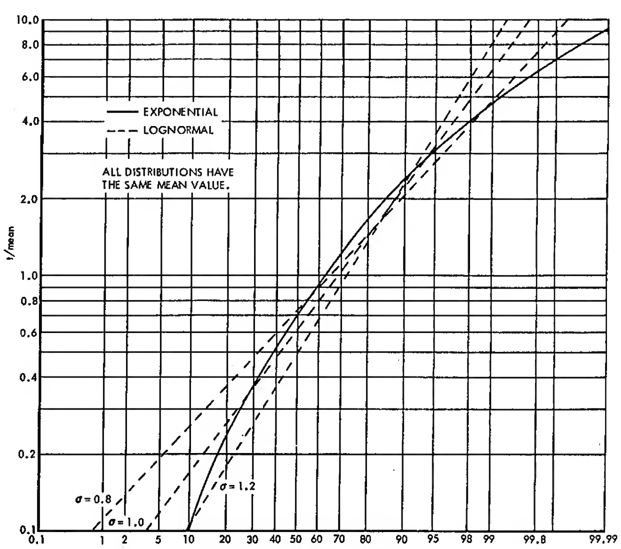


Figure 1. Lognormal and Exponential C.D.F.'s

Initial and Final States

The system under test is assumed to begin in the up state; i.e., assumed to have an initial condition of zero failures.

If samples are not to be censored, the demonstration must also end with no failures in the system. Terminating the demonstration with one or more failures in the system will lead to erroneous results. In effect, if test termination time occurs while failures are in the system, the test must be extended until they have been repaired. This statement implies that the demonstration cannot really be a fixed time test. However, it will be shown that for reasonably long test times, the requirement to repair all failures before test termination is of little concern.

Ergodicity

Classically, availability of a system is the probability that the system is operative at any point in time. This demonstration technique uses accumulated down time as a decision variable. Therefore, the process must be assumed ergodic. That is, the time statistics are assumed equal to the ensemble statistics. This assumption is often made in specifications, for, whenever someone states he expects his system to be usable 95 percent of the time, he is assuming an ergodic process.

The ergodic assumption, together with specified equilibrium availability, allows availability to be demonstrated on a long sample from a single system. In contrast, point availability in the transitional state can only be demonstrated by replications of the same test (on a single or multiple systems) to develop a probability distribution.

Notation

This section summarizes the terms used throughout the remainder of the paper.

- A_s – Specified equilibrium availability
- A – Actual availability
- λ – Average failure rate of the system under test
- μ – Average restoration rate of a single failure in the system
- T – Time duration of the test
- m – Expected down time of the system due to a single failure $m = 1/\mu$
- C – Expected up time of the system $C = 1/\lambda$
- x – Random variable of down time $m = E(x)$
- y – Random variable of the sum of discrete down times $y = \sum_i x_i$
- $I_n(w)$ – Modified Bessel function of order n and argument w
- R – Risk = Pr {failing demonstration}
- E – Expectation operator
- f – The ratio $\mu(1 - A_s)/\lambda$
- $F(\cdot)$ – Cumulative distribution function

State Description

Probably the most significant departure this paper makes from conventional modeling techniques is that of allowing multiple failures. It is usually assumed that once failed, a system cannot fail again. It will be shown that this can be a tenuous assumption for a large system that is to be observed over many up-down cycles. In addition, electronic systems of any size are seldom powered down during trouble-shooting. It is rarely possible to find the problem in a "cold" system. And, so long as power is applied, the system can fail.

Allowing multiple failures leads to realizations such as that shown in Figure 2. Note in the figure that any performance substates within the up state are collected into a single, minimum state. Note also that

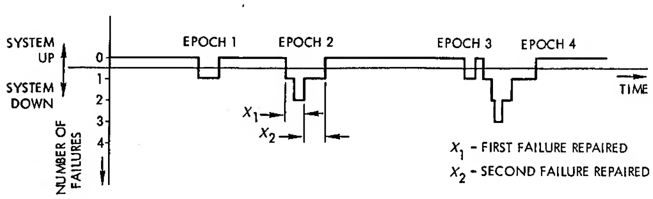


Figure 2. A System State Realization

the state of the system (number of failures) is identical to the state of the repair facility; i.e., the number of failures which are either being repaired or waiting to be repaired. Since the failures are repaired one by one, the system down time for any epoch of failures is simply the sum of the repair times for each failure. In Figure 2, down time for the second epoch is simply $x_1 + x_2$.

The occurrence of failure epochs or multiple failures gives rise to the first indication that conventional approaches to availability demonstrations may produce misleading results. From Equation (A4) in the appendix, the expected duration of system down time for a single epoch, E(d), is $\frac{1}{\mu - \lambda}$, not $1/\mu$. However, for short observations of the system and for $\lambda << \mu$, the approximation of expected system down time, E(d) = $1/\mu$, is usually satisfactory.

A final observation remains to be made before moving on to the risk function. This observation is an extrapolation of the summed repair times to achieve epoch down time previously discussed. Note that system down time for a demonstration is the sum of all the epoch down times. Now each epoch down time is the sum of individual down times. Since none of these times is overlapping, it follows immediately that total down time for a system is the sum of the individual repair times. Further, if down time is all that is of interest, it makes no difference when the failures occur. That is to say, if six failures occurred during a demonstration, it is irrelevant whether they occurred individually, all in a single epoch, or in any combination between these two extremes.

Development of the Risk Function

This section develops an analytical model for producer's demonstration risk; i.e., the probability of failing the demonstration. The development deliberately retains the variables λ and μ so that risk may be assessed directly in terms of these elemental values when evaluating design alternatives. As a result, the development is independent of any preconceived availability formulation.

Success Criterion

Using the assumption of ergodicity, the system under test will be accepted if the total time the system is capable of performing its intended function, divided by total test time, is equal to or greater than the specified availability. Accept if,

$$\frac{\text{(operable time)}}{T} \geq A_s \tag{1}$$

Let

OP = Operable time

D = Total down time

and

$OP + D = T$

Subtracting unity from both sides of Equation (1) and multiplying both sides by $-T$,

$$T - OP \leq (1 - A_s)\, T$$

The acceptance criterion is then, accept if

$$D \leq (1 - A_s)\, T \tag{2}$$

This is the desired result, since down time is a directly measurable quantity in a demonstration.

Derivation

The mutually exclusive and exhaustive events which constitute success can be expressed as a probability statement.

$$\Pr\{\text{Passing test}\} = P_0(T) + P_1(T)\, P\{X \leq (1 - A_s)\, T\} + P_2(T)\, P\{X_1 + X_2 \leq (1 - A_s)\, T\} + \ldots \tag{3}$$

Equation (3) is an infinite sum and

$P_i(T)$ = probability of exactly i failures in test time, T.

$P_i(T)\, P\{\sum_i X_i \leq (1 - A_s)\, T\}$ = joint probability of exactly i failures and the sum of the down times is equal to or less than the allowed down time, $(1 - A_s)\, T$.

Now,

R = Risk = 1 - Pr{passing test}

and

$$R = 1 - P_0(T) - P_i(T)\, P\{X \leq (1 - A_s)\, T\} - \ldots \tag{4}$$

The first two terms in Equation (4) are

$$P\{\text{one or more failures}\} = 1 - P_0(T) = \sum_{n=1}^{\infty} P_n(T)$$

Rearranging Equation (4)

$$R = \sum_{n=1}^{\infty} P_n(T) - P_1(T)\left[1 - P\{X > (1 - A_s)\, T\}\right] - P_2(T)\left[1 - P\{X_1 + X_2 > (1 - A_s)\, T\}\right] - \ldots$$

Carrying out the obvious substitutions

$$R = P_1(T)\, P\{X > (1 - A_s)\, T\} + P_2(T)\, P\{X_1 + X_2 > (1 - A_s)\, T\} + \ldots \tag{5}$$

From the assumption of exponentially distributed times between failures, the $P_i(T)$ are Poisson,

$$P_i(T) = \frac{e^{-\lambda T}(\lambda T)^i}{i!} \tag{6}$$

From the assumption of exponentially distributed restoration times, the sum of the down time variates is Gamma distributed; i.e., the α-fold convolution of the exponential density. (Note that the X_i are independent samples from the same population.) In particular,

$$P\{\sum_{\alpha} X_{\alpha} > (1 - A_s)\, T\} = 1 - F_{\alpha}(y)$$

$$= \int_{(1-A_s)T}^{\infty} \frac{\mu^{\alpha}}{(\alpha - 1)!}\, y^{\alpha-1}\, e^{-y\mu}\, dy$$

Making use of the identity which relates the Gamma c.d.f. to the Poisson distribution,

$$1 - F_{\alpha}(y) = \sum_{k=0}^{\alpha-1} \frac{(\mu z)^k e^{\mu z}}{k!} \qquad (7)$$

where

$$z = (1 - A_s)\, T \quad ,$$

an expression for exceeding a fixed down time requirement can now be written as a double summation. Using Equations (6) and (7) in Equation (5) and factoring common terms in the series, risk can be expressed

$$R = e^{-(\lambda T + \mu z)} \sum_{k-1}^{\infty} \sum_{n=0}^{k-1} \frac{(\lambda T)^k}{k!} \frac{(\mu z)^n}{n!} \quad . \qquad (8)$$

Note that the risk for $T = 0$ is zero.

Equation (8) is a good interim result for numerical solution and was used to plot the risk graph in the section on **The Risk Graph and Its Use.** It will, however, be necessary to examine the asymptotic behavior of the function. This will necessitate altering its form in the next section on **Asymptotic Behavior** in order to achieve further results.

Before departing the discussion of the risk function, implications of requiring the test to terminate with no failures should be addressed. This is the topic of the next subsection.

Test Termination Criteria

It has been indicated that, if one or more failures exist in the system under test when the fixed test time expires, the test must continue until these failures are repaired (and any subsequent failures which occur while repair is being effected). Otherwise, termination will result in sample censoring since

$$\frac{a}{T} \neq \frac{a - b}{T - b} \quad ,$$

where a is the total down time if the test were allowed to continue and b is the amount by which the repair time was truncated.

There is, then, a finite probability that the test will last longer than T and this would influence the risk. Fortunately, this probability is quite small. Equation (A2) of the appendix indicates that the probability of exactly one failure in the system at T, after the process has been operative for more than four hours, is $(1 - \lambda/\mu)(\lambda/\mu)$.* This will typically be a very small value. If, in addition, $T > 10m$, the influence of this failure, should it exist, will be quite small. To assess the likelihood of multiple failures present in the system at T, it can be stated that

$$P_n = (1 - \lambda/\mu)(\lambda/\mu)^n$$

where n is the number of failures. Note that the probability of zero failures is $1 - (\lambda/\mu)$.

Asymptotic Behavior

This section presents the limiting risk values as T goes to infinity. Equation (8) will first be placed in a form for which limiting behavior is recognizable and the limits for three values of availability (in terms of λ, μ) will be examined.

Reformulation of the Risk Function

For convenience of notation let

$$a = \lambda T$$

$$b = \mu z = \mu(1 - A_s)\, T \quad .$$

Equation (8) can then be expanded

$$R = e^{-(a+b)} \left\{ a + \frac{a^2}{2!}(1 + b) + \frac{a^3}{3!}\left(1 + b + \frac{b^2}{2!}\right) + \ldots \right\} \quad .$$

Rearranging terms

$$R = e^{-(a+b)} \left\{ \left(a + \frac{a^2}{2!} + \frac{a^3}{3!} + \frac{a^4}{4!} + \ldots\right) + a\left(\frac{ab}{2!} + \frac{(ab)^2}{3!\,2!} + \frac{(ab)^3}{4!\,3!} + \ldots\right) + a^2(\ldots) + \ldots \right\} \quad .$$

The first term in the braces is $e^a - 1$. The remaining terms will require further manipulation. If unity is added and subtracted from the group of terms forming the coefficient of a^1, 1/2 added to and subtracted from the a^2 coefficient and, in general, add and subtract $1/n!$ from the a^n coefficient, the terms form a sum of Bessel functions plus a power series in a. This power series is $-(e^a - 1)$ which cancels the original exponent and, finally,

$$R = e^{-(a+b)} \sum_{n=1}^{\infty} \left(\sqrt{\frac{a}{b}}\right)^n I_n\left(2\sqrt{ab}\right) \quad . \qquad (9)$$

Limiting Risk when $a = b \Rightarrow \lambda = \mu\,(1 - A_s)$

From the identity

$$e^w = I_0(w) + 2 \sum_{n=1}^{\infty} I_n(w) \quad ,$$

it follows that

$$\sum_{n=1}^{\infty} I_n(w) = 1/2\, e^w - 1/2\, I_0(w) \quad .$$

Letting $a = b = c$ and substituting this result into Equation (9),

$$R = e^{-2c}\left[1/2\, e^{2c} - 1/2\, I_0(2c)\right]$$

$$R = 1/2 - 1/2\, e^{-2c}\, I_0(2c) \quad .$$

Recalling that c is a function of test time,

$$\lim_{T \to \infty} R = 1/2$$

*From the appendix, the four-hour interval represents the maximum time required for most electronic systems to reach equilibrium.

since

$$\lim_{w\to\infty} e^{-w} I(w) = 0$$

Limiting Risk when $a<b \Rightarrow \lambda<\mu(1-A_s)$

Substituting the asymptotic Bessel expression

$$I_n(w) \sim \frac{e^w}{\sqrt{2\pi w}} \quad \text{as } w\to\infty$$

into Equation (9) and noting the expression is independent of n,

$$R = e^{-(a+b)} \frac{\exp(2\sqrt{ab})}{\sqrt{2\pi(2\sqrt{ab})}} \left[\sqrt{\frac{a}{b}} + \left(\sqrt{\frac{a}{b}}\right)^2 + \left(\sqrt{\frac{a}{b}}\right)^3 + \ldots\right]$$

Since $a<b$ and a, b are greater than zero, the sum goes to

$$\frac{1}{1-\sqrt{\frac{a}{b}}} - 1 = K$$

which is a constant greater than zero since a/b is time invariant. Then

$$R = \frac{\exp\left[2\sqrt{ab} - (a+b)\right]}{\sqrt{2\pi(2\sqrt{ab})}} + K.$$

Note that the denominator is proportional to $\sqrt{T}$ and note further that $a+b$ is always larger than $2\sqrt{ab}$ if $a\neq b$. Then,

$$\lim_{T\to\infty} R = 0$$

Limiting Risk when $a>b \Rightarrow \lambda>\mu(1-A_s)$

Substituting the Bessel generating function

$$\exp\left[\frac{w}{2}\left(y+\frac{1}{y}\right)\right] = \sum_{k=0}^{\infty} y^k I_k(w) + \sum_{k=1}^{\infty} y^{-k} I_k(w)$$

into Equation (9)

$$R = e^{-(a+b)}\left\{\exp\left[\sqrt{ab}\left(\sqrt{\frac{a}{b}}+\sqrt{\frac{b}{a}}\right)\right] - \sum_{k=1}^{\infty}\left(\sqrt{\frac{a}{b}}\right)^{-k} I_k\left(2\sqrt{ab}\right) - I_0\left(2\sqrt{ab}\right)\right\}$$

It has already been shown that the last two terms of this equation, when multiplied by $e^{-(a+b)}$, go to zero in the limit. This makes use of the facts that $a\neq b$ and $(\sqrt{a/b})^{-k}$ is less than unity. Now, the first term in braces is e^{a+b}. Therefore

$$\lim_{T\to\infty} R = 1$$

Summary of Results

	Limiting Value of R
$\lambda = \mu(1-A_s) \Rightarrow A_s = (1-\lambda/\mu)$	1/2
$\lambda<\mu(1-A_s) \Rightarrow A_s < (1-\lambda/\mu)$	0
$\lambda>\mu(1-A_s) \Rightarrow A_s > (1-\lambda/\mu)$	1

(10)

Put still another way, if the system is designed such that $\lambda/\mu > (1-A_s)$, the risk on a very long test will tend to unity. If the sense of the inequality is reversed, the risk on a very long test will tend to zero.

The equality relationship above is significant. It states in effect that $A = 1 - \lambda/\mu$, not that

$$A = \frac{c}{m+c} = \frac{\mu}{\lambda+\mu} \tag{11}$$

which is the expression most commonly encountered. The difference in the two expressions can be more than academic for extended availability demonstrations, leasing contracts containing incentives on availability, or life-cycle costing evaluations. Consider the case where a designer is faced with meeting a specified availability, A_s. The first step he must take is to convert this requirement into the direct design parameter λ/μ. Equation (11) can be rewritten to express this relationship.

$$A_s = A = \frac{1}{1+\frac{\lambda}{\mu}}$$

Solving this equation for the equivalent design constraint,

$$\frac{\lambda}{\mu} = \frac{1-A_s}{A_s}$$

Now, A_s exists on the closed interval [0, 1]. In practice, however, it is reasonable to assume $A_s < 1$. Thus, the constraining value for λ/μ arrived at in this fashion must be greater than $(1-A_s)$. It follows that a system which exactly met an apparently good constraint derived from Equation (11) would realize a risk of unity over a very long demonstration. Note that the difficulty arises, not from the fact that the two expressions yield a slightly different value of availability for the same system, but from the use of Equation (11) to identify design constraints.

The Risk Graph and Its Use

This section describes the risk graph developed from Equation (8). It is recognized that Equation (8) must be solved on a computer and that this would limit its practicality, especially during conceptual design when strategies must be formulated on rough inputs. As a result, it is imperative that a versatile plot be formulated such that graphical solutions can be achieved for a wide range of problems. The resulting graph is shown in Figure 3. Logscales have been used to extend the range. Users will seldom desire risk accuracies any better than plus or minus two points, but they will typically have at least two place accuracy on λ and μ. With this in mind, the inverse risk function has been plotted in the form of iso-risk lines. The straight, diagonal line bisecting the graph divides the regions which are above specification and below specification: the top, left being the above specification region.

Reading the Graph

The abscissa of the graph is normalized on system mean restoration time. The scale is then in multiples of this quantity.

To determine the risk expected to be incurred on an availability demonstration, one need know test time (T), the mean system failure rate (λ), the specified availability (A_s) and the mean system restoration rate (μ). These quantities determine a unique value for risk which may be interpolated from the plotted risk curves. Any points lying to the left of or above the 1 percent risk curve yield a risk less than one percent. Likewise, points lying to the right of or below the 95 percent curve yield a risk greater than 95 percent.

Note that T appears on both axes. Thus, for given values of λ, μ

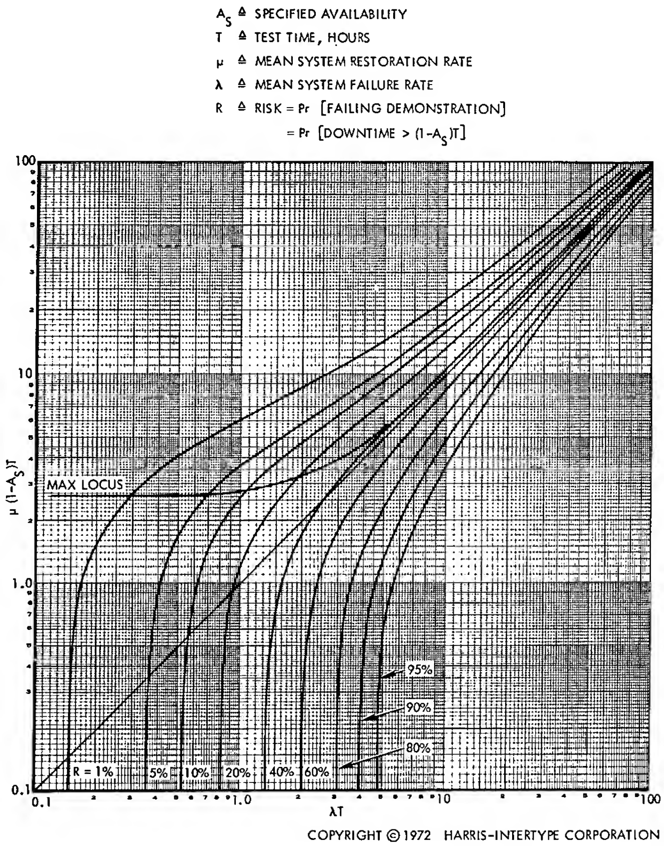


Figure 3. Availability Risk Graph

and A_S, a 45-degree line will plot the locus of risk as a function of time. This is demonstrated in Example 3 below. The curve labeled "max locus" will also be explained in Example 3.

Example 1

It has been determined that a system to undergo a demonstration has the following availability parameters.

$\lambda = 5 \times 10^{-3}$ failures per hour

$\mu = 0.4$ restorations per hour

The system is to be tested for 160 hours during which time the total down time cannot exceed 4 hours. What is the risk?

Note that the expression $(1 - A_S)\,T$ is actually the maximum down time allowed during the test. Thus

$(1 - A_S)\,T = 4.0$

Now, $\mu(1 - A_S)\,T = 1.6$ and $\lambda T = 0.8$. These points are plotted in Figure 4 and yield a risk of approximately 12 percent.

Example 2

A system currently in the conceptual phase is to undergo an availability demonstration as part of Category III Testing. An availability of 95 percent is specified and the test lasts 1,000 hours.

Initial estimates yield the following bounds on the availability parameters.

$0.0125 \leq \lambda \leq 0.0167$

$0.66 \leq \mu \leq 1.0$

What is the expected risk and where should efforts be concentrated to gain the greatest risk reduction?

Now,

$12.5 \leq \lambda T \leq 16.7$

$20 \leq \mu(1 - A_S)\,T \leq 30$

These points are plotted in Figure 4 as an operating region. The worst-case risk is 30 percent and the most optimistic is less than 1 percent. At this particular position on the graph, the risk lines come very close to forming 45° angles to the rectangular operating region. However, the slope is still somewhat less than unity and efforts to increase μ will reduce risk slightly more than equal efforts at decreasing λ.

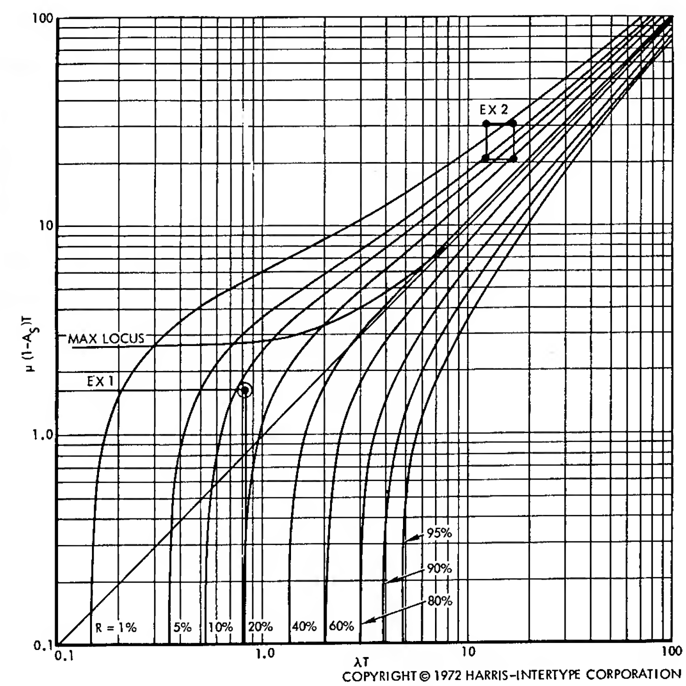


Figure 4. Risk Plots for Examples 1 and 2

Note that if the operating region had fallen below the knee of the curves, the decision is obvious. Increasing μ will produce very little reduction, but decreasing λ will have a radical effect.

Example 3

A system in the conceptual phase is to be designed to meet an availability of 99 percent. The system has a failure rate of 6×10^{-3}. Three different maintainability approaches are being considered. The approaches have the following restoration rates.

Approach 1, $\mu = 0.5$

Approach 2, $\mu = 0.75$

Approach 3, $\mu = 1.2$

(a) Do any of the approaches fall below specification? (b) What test time should be used? (c) What is the maximum risk encountered?

This example involves construction of a time structure on the risk graph. Recall that time increases along 45-degree lines. To form a time line, it is only necessary to form the ratio $f = \mu(1 - A_S)/\lambda$ which establishes the value of $\mu(1 - A_S)\,T$ per unit value of λT. Convenient points are then picked along the λT axis (usually 1.0 and 10.0). The corresponding value of $\mu(1 - A_S)\,T$ is known by the ratio and the points plotted. The points are then connected by a straight line which indicates risk as a function of time for a given set of parameters.

To continue, the ratios for the three approaches will now be formed.

1) $\mu = 0.5$

$$f = \frac{\mu(1 - A_S)}{\lambda} = 0.834$$

2) $\mu = 0.75$

$$f = 1.25$$

3) $\mu = 1.2$

$$f = 2.0$$

Question (a) can be answered immediately. Since f for Approach (1) is less than unity, this approach falls below specification. The remaining approaches exceed specification.

Though below specification, it is instructive to carry Approach (1) through the example and construct time lines for all three approaches. Using f from Approach (1), locate λT equal to unity and 10.0. Corresponding values of $\mu(1 - A_S)\,T$ are then $1.0f$ and $10.0f$. These points are then connected by a straight line as shown in Figure 5. The procedure is repeated for the remaining two approaches.

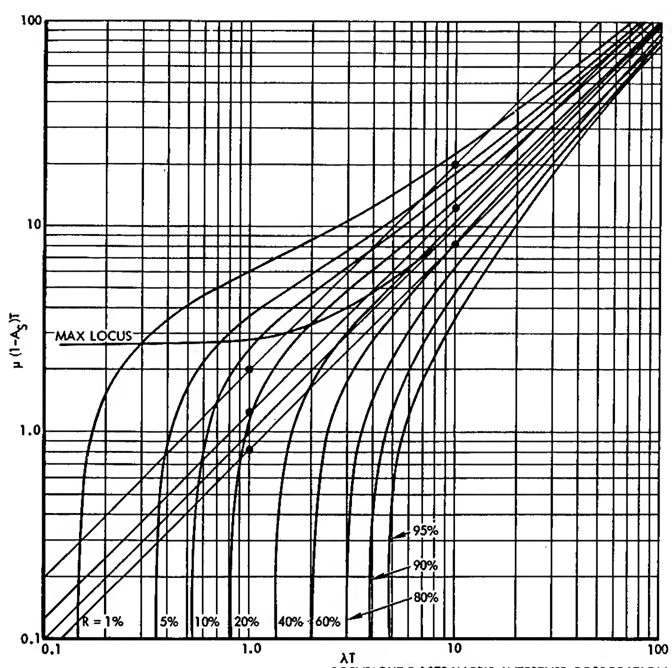


Figure 5. Time Structure Lines Overlaying Risk for Example 3

It is difficult to view the risk when plotted in this fashion. Fortunately, it is an easy matter to transform the information into a more conventional format. Replot the abscissa on another piece of graph paper. Then read the risk values on the time line corresponding to λT. Plot the values. The results are shown in Figure 6. The dashed curved labeled $f = 1.0$ is discussed below.

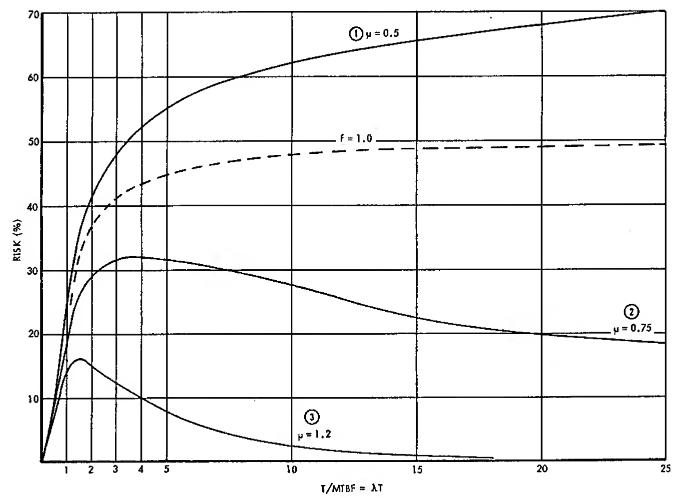


Figure 6. Risk as a Function of Time for Example 3

The plots of risk versus time in Figure 6 are quite revealing. Note that risk for Approach (1) increases dramatically. Unlike the other two curves, risk for Approach (1) will continue to increase, becoming asymptotic to 100 percent.

The two compliant approaches exhibit curves which quickly peak and gradually diminish. These curves will asymptotically approach zero. Note that all the curves begin at the origin. The curve labeled $f = 1.0$ is that formed when $\lambda = \mu(1 - A_S)$. Curves lying to the left of this plot will break to a risk of 100 percent. Curves to the right will break to zero. Note that the dashed curve is asymptotic to a risk of 50 percent.

Question (b) can now be addressed. For Approach (2), if one is willing to accept a 30 percent risk, the demonstration time would have to exceed 7.5 MTBF. If a 20 percent risk is considered maximum, test time must be greater than 20 MTBF. These times are 1,250 and 3,320 hours, respectively.

Risk for Approach (3) is always less than 18 percent. If this risk is acceptable, a very short time can be used (but greater than 10 $1/\mu$). If for some reason the test time had to exceed one MTBF, it would be desirable to test for at least four MTBF's to avoid the peak risk below this value. In any event, any additional cost to implement Approach (3) might well offset the cost incurred for the long test times and comparatively high risk encountered in Approach (2).

The discussion of peak risk leads to the answer for Question (c). From Figure 6, maximum risk does not really apply to Approach (1). Approaches (2) and (3) have maximum risks of 32 percent and 17 percent, respectively. An important point to note here is that Figure 6 did not have to be plotted to determine these results. The line on the Availability Risk Graph, Figure 3, labeled "max locus" identifies the points immediately. It is the locus of the maxima for all solutions which do not represent ever-increasing risk; i.e., those within specification. The intersection of a time line with max locus is the maximum risk for that line.

Conclusions

The availability demonstration technique presented in this paper is believed to be a viable approach. In addition, the risk assessment formulation, together with the developed graphical solution capability,

yield a comprehensive and tractable methodology for determining availability strategies and performing tradeoffs in early design stages.

The following pertinent points may also be concluded:

- If $\lambda > \mu(1 - A_s)$, risk will tend to unity over a long demonstration.
- If $\lambda < \mu(1 - A_s)$, risk will tend to zero over a long demonstration.
- Using system down time as a decision variable allows the natural introduction of queued maintenance actions into availability demonstration risk.
- Assuming a large system may not fail while it is being repaired can lead to understatements about risk for extended demonstrations. This is evidenced by using the formulation $A_s = \mu/(\lambda + \mu) = 1/(1 + \lambda/\mu)$ to solve for an equivalent specified constraint in terms of λ/μ. Demonstrating a system which just met this constraint would lead to a risk of unity for a very long test. Thus, constraints for λ/μ derived in this fashion are obviously erroneous. Solving for constraints using $A_s = 1 - (\lambda/\mu)$ will not produce these results. This leads to the next conclusion.
- A more appropriate expression for availability might be $A = 1 - (\lambda/\mu)$.
- Under the stated constraints, it is very likely that scheduled test time, T, will increase less than 10 percent due to failures existing at scheduled termination time. Thus, the derived risk is subject to only minimal change during the demonstration.
- The significant body of knowledge surrounding the Queue M/M/1 is directly applicable to availability modeling and formulation (see the appendix).

Appendix
Relationships to the Queue M/M/1

Properties of the Queue M/M/1 will be summarized in this appendix and related to the availability risk analysis. The material has been drawn from Saaty,[4] Prabhu,[5] and Cox.[6] Results of the appendix assume $\lambda < \mu$.

Description of the Queue M/M/1

This section describes the general operation and states of the queue as well as the notation used in the appendix.

Operation

A queueing system consists of a server, a waiting line, and a calling population. The server performs some operation on or for "customers." In this case the "customers" are failures which must be repaired by a single repair team or facility. The team can repair but one failure at a time. The waiting line consists of failures awaiting their turn to be repaired. A waiting line can take on many forms. Here, the length of the line is not restricted and may get infinitely long. Since the system under test is down so long as a failure exists, there will be no concern with whether the service discipline is first-come-first-served or not. The calling population is the number of "customers" which may be interested in securing the service offered by the server. Here, it is all the failures which may occur in the system under test. It is safe to assume that this population is infinite in size.

A single server queue with infinite calling population, infinite queue size, and arbitrary service discipline has just been described. To complete this description, the arrival and service distribution must be identified. Since the M/M/1 queue is being used as a model, inter-arrival time of failures is assumed exponentially distributed. That is, the time between single failures behaves as the exponential distribution. Also, the restoration (service) time is assumed exponentially distributed.

States

State of the queue will be defined as the number of customers in the waiting line and the service facility combined. Thus, a state of three means that there is one failure being repaired (serviced) plus two waiting. Note that this is exactly the state description given in the risk assessment. When the queue is in state zero, there are no failures being repaired or waiting and the system under test is in the up state.

Notation

The following notation is used throughout the appendix.

$I_n(w)$, T, λ, μ, A_s – Same as indicated in the body of the paper.

$P_n(t)$ – Probability that the queue is in state n at time t under stated initial conditions.

$P(n)$ – Probability that the queue occupies state n under equilibrium conditions.

t – Time measured from the point at which initial conditions existed.

State Zero Probability

Availability is the probability that a system is up at any point in time. Based on the argument above, the probability of the queue being in state zero is then identically equal to availability. Under the defined initial conditions, viz.,

$$P_0(0) = 1$$

$$P_n(0) = 0, n = 1, 2, 3 \ldots\ldots$$

the probability of being in state zero is

$$A(t) = P_0(t) = e^{-(\lambda + \mu)t}\{I_0(2t\sqrt{\lambda\mu}) + \sqrt{\mu/\lambda}\, I_1(2t\sqrt{\lambda\mu}) + (1 - \lambda/\mu)\sum_{k=2}^{\infty} I_k(2t\sqrt{\lambda\mu})\} \quad \text{(A1)}$$

Inspection will show that $P_0(0)$ is unity which satisfies the initial conditions. As for final conditions, $P_0(\infty)$, arguments similar to those used in the analysis of asymptotic behavior of the risk function will show that $P_0(t)$ is asymptotic to $1 - (\lambda/\mu)$. This result is the equilibrium solution for availability. From a practical viewpoint, $P_0(t)$ will complete 90 percent of its transition to final value within 4 hours for most electronic systems.

Solving the M/M/1 equilibrium state equations, the following general result is obtained.

$$P(n) = (1 - \lambda/\mu)(\lambda/\mu)^n; n, \text{int.}, \geq 0 \qquad \sum_n P(n) = 1 \quad \text{(A2)}$$

Equation (A2) is then the probability of finding the system under test with n failures at some point in time.

Busy and Idle Periods

The busy period of a queue begins when a single customer arrives at an idle server and ends when the server next becomes idle. Busy

period takes on significance when customers may arrive while one is being served. The server is then performing operations on these customers in immediate succession until all are served. The mean down time for the system under test must then be greater than $1/\mu$ since this is the time to repair a single, isolated failure.

The density function of the busy period is

$$b(\tau) = (\tau\sqrt{\mu/\lambda})^{-1} e^{-(\lambda+\mu)\tau} I_1 (2\tau\sqrt{\lambda\mu})\, d\tau \qquad . \qquad (A3)$$

The expected duration of a busy period or down time for failure epochs of the system under test is

$$E(d) = \frac{1}{\mu - \lambda} \qquad . \qquad (A4)$$

That is, when the system under test fails, it is expected to be down for $1/(\mu - \lambda)$.

The idle period of a queue begins when the server first becomes idle and ends when the next single customer arrives. Since exponential interarrival times are assumed, the density function of the idle period is simply the exponential density with expectation $1/\lambda$. The mean up time for the system under test is then

$$E(U) = 1/\lambda \qquad . \qquad (A5)$$

It is instructive to use Equations (A4) and (A5) in a familiar relation to derive equilibrium availability by another method. Let

$$A = \frac{E(U)}{E(U) + E(d)}$$

then

$$A = \frac{\lambda(\mu - \lambda)}{\lambda\mu}$$

and,

$$A = 1 - \lambda/\mu \qquad . \qquad (A6)$$

Equation (A6) yields the same results as Equation (A2) and the equality expression of Relations (10).

Queueing Interpretation of the Risk Function

Recall from the risk function that the quantity $(1 - A_s)T$ is the critical aggregate down time. Alternatively, it is the total time the system (and the queue) spends outside of state zero. $(1 - A_s)T$ is then the aggregate time the queue is busy. The risk may then be interpreted as the probability that the queue stays busy greater than some aggregate time D over a total period T, $D < T$, given the queue was idle initially. The quantity $(1 - A_s)T$ is then replaced by D. Entering Figure 3 with these values will yield the probability graphically.

References

1. M. A. McGregor, "Approximation Formulas for Reliability with Repair," **IEEE Transactions on Reliability**, December 1963, pp.64–92.

2. S. J. Einhorn, "Reliability Predictions for Repairable Redundant Systems," **Proceedings of the IEEE**, Vol. 51, No. 2, February 1963, pp.312–317.

3. A. S. Goldman and T. B. Slattery, **Maintainability: A Major Element of Systems Effectiveness,** John Wiley and Sons, 1964.

4. T. L. Saaty, **Elements of Queueing Theory**, McGraw-Hill, 1961.

5. N. U. Prabhu, **Queues and Inventories,** John Wiley and Sons, 1965.

6. D. R. Cox and H. D. Miller, **The Theory of Stochastic Processes**, John Wiley and Sons, 1968.

AN ALGORITHM FOR OPTIMIZING THE SAFETY EFFECTIVENESS OF PARALLEL ALARM SYSTEMS

INDEX SERIAL NUMBER – 1128

BY

Igor Bazovsky, Jr.

Igor Bazovsky and Associates
Sherman Oaks, California

Abstract

A computational algorithm is given to determine the number of alarm systems which minimize the expected total Life Cycle Costs in the following situation. The alarm systems monitor a critical component and when they give an in-time alarm then the costs of failure are relatively minor, but when they fail to give an alarm when called for (i.e. either no alarm or a tardy alarm) then the costs of failure are severe. The paper shows that:

i. under general circumstances the expected total life cycle cost is a unimodal function (and hence whose minimum can be achieved by a Fibonacci search in at most F_m evaluations).

ii. under some common conditions an explicit expression for the optimum is available.

Furthermore a sufficient condition is given for ensuring that the most advantageous ordering is used in selecting the alarm systems. In this context the notion of best ordering is found useful.

Scenario

Consider an alarm system which has several independent components sensing a critical parameter, and suppose that each one has the capability of sounding the alarm or warning by itself when the critical parameter exceeds its redline limits. Now, for a component to sound the alarm it must be in a non-failed state and it must react properly to the sensing of the out-of-specification signal. Furthermore, the minimum reaction time of the ones which do function properly must be small enough to give the operator (man or machine) sufficient time to react in a fail safe manner. The situation is represented schematically in figure 1.

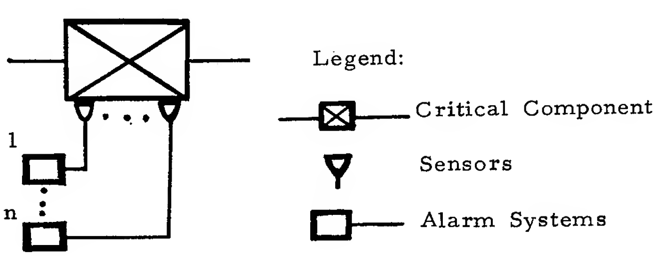


Figure 1: n Alarm Systems in Parallel

The only way an accident can occur is for the critical component to fail and none of the alarm systems functions in time. The probability that one of the alarm systems does not function in time is the complement of the probability that it has not failed in the quiescent mode, that it switches properly in response to a call for alarm, and that it gives an in-time alarm. Thus

$$P(A)=P(\text{Accident}) = \int_0^T \left\{ \prod_{i=1}^{n} (1-E_i(t)) \right\} f(t)dt \qquad (1)$$

where, $E_i(t)$ = Instantaneous Effectiveness

$= r_i R_i(t) H(a(t))$.

r_i = switching reliability of i th. alarm system.

$R_i(t)$ = quiescent mode reliability of i th alarm system.

$a(t)$ = maximum permissable delay time in responding to a call for alarm.

$H_i(x)$ = is the density function of the reaction time of the i th. alarm system.

T = the mission time.

$f(t)$ = the pdf of time-to-occurence of critical parameter in an out-of-specification condition.

For later work the following conditions are needed.

Condition A: There is a number s, $0<s<1$, such that:

Instantaneous Effectiveness = $E_i(t) \geq s$

for all $t<T$ and for all i. Condition A assures that the alarm systems being put into the system do not get below a certain minimum level of effectiveness.

Condition B: The alarm systems will be said to satisfy condition B if they have been ordered so that

$$E_i(t) \geq E_{i+1}(t)$$

for all $t<T$ and for all integers i.

Remark on Condition B: Clearly condition B is not universally obtainable. However under some common conditions one can arrange the alarm systems to satisfy it. Some situations in which condition B is attainable are:

i. All the alarm systems are similar and have identical properties at all possible mounting locations.

ii. The quiescent failure rates λ_i of the alarm systems are constants, and the rest of the characteristics are identical. Then ordering the systems so that $\lambda_i < \lambda_{i+1}$ assures condition B.

iii. If $a(t) = a$ for all $t<T$ and if the alarm systems can be so arranged so that

$$r_{i+1}H_{i+1}(a) \leq r_i H_i(a),$$

then the equality of the quiescent failure rates is sufficient to assure condition B.

Of the above, situation ii is most practical because it accomodates the case of identical alarm systems whose quiescent reliability is determined by the ambient environments at each individual location.

Cost Model

Suppose that the addition of the n th alarm system necessitates an additional cost of $c_1(n)$ dollars. (This includes procurement costs,

installation costs, maintenance costs, power expenditure costs, weight and space penalty costs, and the loss due to profit losses resulting from filling up the space by alarm systems instead of by profit making devices.) In general $c_1(n)$ will not be constant. Some factors will tend to raise it (weight, space and power saturation effects) and others will tend to lower it (combined maintenance activity, lower procurement, proration of user training etc.). Denote by c_2 the average cost per non-accident failure and by c_3 the average cost per accident failure. Assuming independence and that all the alarm systems are replaced after each critical component failure, the general expression for the expected total cost is,

$$E\left[C \;\middle|\; n\right] = \sum_{i=1}^{n} c_1(i) + c_2 E\left[NA\right] + c_3 E[A]$$

$$= \sum_{i=1}^{n} c_1(i) + \left\{ c_2 p_n + c_3 (1-p_n) \right\} E\left[N\right]$$

where NA = number of non-accident failures
A = number of accident failures
N = number of failures
p_n = probability of no accident when n alarm systems are used and a failure is known to have occurred.
$E\left[C\middle|n\right]$=Expected cost given n alarm systems.

The optimum number n_o which minimizes $E\left[C\middle|n\right]$ is conveniently studied through the quantity,

$$\Delta(n) = E\left[C\middle|n+1\right] - E\left[C\middle|n\right]$$
$$= c_1(n+1) - b(p_{n+1} - p_n) \qquad (3)$$

where $b = (c_3 - c_2)E\left[N\right]$ is positive when and only when the average cost of an accident (i.e. a non alarmed failure) is greater than the average cost of a non-accident failure.

Results

The results can be summarized in the following general procedure and theorems.

General Procedure

If conditions A and B are satisfied and if $c_1(n)$ is non decreasing, guess at a value n_1 for n_o and evaluate $\Delta(n_1)$. If $\Delta(n_1) \geq 0$, then $n_o \leq n_1$. If $\Delta(n_1) < 0$, then $n_o > n_1$. In the latter case, continue picking $n_2, n_3 \ldots (n_1 < n_2 < n_3 \ldots)$, until $\Delta(n_i) \leq 0$. If n_i is the first integer found for which $\Delta^1(n_i) \leq 0$, then $n_{i-1} < n_0 < n_i$. According to theorem 2 below and reference 1, the optimum search procedure is Fibonacci Search and the maximum number of numerical evaluations, once n_{i-1} and n_i have been found is the $(n_i - n_{i-1})$ th. $=m^{th}$ Fibonacci number,

$$F_m = \left(\frac{1+\sqrt{5}}{\sqrt{5}}\right)\left(\frac{1+\sqrt{5}}{2}\right)^m + \left(\frac{\sqrt{5}-1}{\sqrt{5}}\right)\left(\frac{\sqrt{5}-1}{2}\right)^m$$

Theorems 1 through 3 provide guidelines for choosing n_1. Corollary 4 provides an explicit solution for a common case.

Qualitative Results on the Nature of n_o

Theorem 1: For a general non-decreasing real world cost function $c_1(n)$, $n_o \leq n_1$ where n_1 is the largest integer not larger than the largest root of $c_1(x+1) - b = 0$.

Theorem 2: If $c_1(x+1)$ is a real world, non-decreasing function, and if condition B is satisfied, then $\Delta(x) = 0$ has at most one root, x_0, and $x_0 < \infty$. If $\Delta(1) > 0$ then $\Delta(x) = 0$ has no root and $n_0 = \Omega$. If $\Delta(1) < 0$, then $\Delta(x)$ has exactly one root.

Theorem 3: If conditions A and B are satisfied, if $c_1(n)$ is a non decreasing function of n, if all the alarm systems have identical properties, and their quiescent failure rate is zero, if $a(t) = a$ for all $t < T$ and if for any given positive constant c, y(c) is defined as,

$$y(c) = \frac{\ln(c) - \ln\big(rbH(a)\big)}{\ln(1 - rH(a))}$$

and n(c) is defined as the largest integer not greater than y(c), then,

$$n(c) > n_o \geq c_1^{-1}(c), \text{ if } c_1(n(c)+1) > c$$
$$n_o = n(c) \quad , \text{ if } c_1(n(c)+1) = c$$
$$n_o \geq n(c) \quad , \text{ if } c_1(n(c)+1) < c$$

where $c_1^{-1}(c) = \left\{ x \;\middle|\; c_1(x) = c \right\}$.

Corollary 4: Under the conditions of theorem 3, if,

$$c_1(n+1) = \begin{cases} c, & \text{for } n < n_1 \\ \infty, & \text{for } n \geq n_1 \end{cases}$$

then,

$$n_o = \begin{cases} 0, & \text{if } \Delta(1) > 0. \\ n(y) \text{ or } n(y)+1, & \text{if } \Delta(1) < 0 \text{ and } y(c) \leq n_1 \\ n_1 & , \Delta(1) < 0 \text{ and } y(c) > n_1 \end{cases}$$

Remark: Corollary 4 corresponds to the conventional linear constraints. E.g. chapter 6, section 2, Optimal Allocation of Redundency Subject to Constraints (2).

Best Ordering

When conditions A and B are satisfied by a given set of alarm systems, theorem 3 or corollary 4 given above can be used to find n_o. However n_o is a function, $n_o(\Omega)$, of the particular ordering Ω used.

Definition: A **best** ordering Ω_o is one for which $n_o(\Omega_o) \leq n_o(\Omega)$ for all orderings Ω.

A given set of alarm systems can have several best orderings. Theorem 5 below shows that under some circumstances a completely dominated ordering is essentially unique and in others is easy to find from basic principles.

Theorem 5: If the cost function $c_1(n)$ is independent of the particular ordering used, then when condition B is satisfied by an ordering $\mathbf{\Omega}_o$, $\mathbf{\Omega}_o$ is a best ordering.

Remark: Site Selection. The condition on $c_1(n)$ required by theorem 5 will be obtained in the important case ii given in: "Remark on Condition B" above, when, as mentioned previously, the only reason for the difference in quiescent reliability is that the ambient environment varies in each location. In this case the theory is useful for selecting mounting locations.

A refinement of the whole theory is possible by considering the case when alarm systems themselves are destroyed whenever an accident occurs. In that case equation 2 needs a slight revision to account for the different costs. Note that as it stands, equation 2, assumes that all the alarm systems are replaced at every failure. When their quiescent reliability is exponential, then due to the memory-less nature of the exponential distribution, this is equivalent to just replacing the failed ones. But when this special condition does not hold then the assumption of total replacement of all alarm systems after each failure is necessary for the development of the theory herein.

Number of Accidents PDF

Once the optimum number of alarm systems have been determined, the distribution of the accident costs becomes of interest. This will depend on the number of missions, on operating time etc. The following considers the two common situations of: fixed number of attempted missions, and fixed number of required successes.

Fixed Number of Attempted Missions

This is the case of the Binomial distribution if either constant failure rates or renewal after each attempt is assumed. In that instance:

$$P(\text{ k accidents}) = \binom{N}{k} p^k (1-p)^{N-k}$$

where p = P(accident) as given in equation 1.

Fixed Number of Required Successfull Missions

If again the condition of renewal after each attempt (successful or not) or of constant failure rates with replacement only of failed items is imposed, then this is the case of the Negative Binomial pdf. That the number of unsuccessful attempts is Negative Binomial is well known. However the number of accidents before the M th. success happens to be that also. This follows from basic principles because all non-accident failures can be neglected. Thus,

$$P(\text{success} \mid \text{success or accident}) = \frac{R_c}{R_c+P(A)},$$

$$P(\text{accident} \mid \text{success or accident}) = \frac{P(A)}{R_c + P(A)}$$

where R_c = Mission Reliability of the critical component, and P(A) =P (accident). Thus:

Theorem6: P(k accidents) =

$$\binom{M+k-1}{k}\left(\frac{R_c}{R_c + P(A)}\right)^M \left(\frac{P(A)}{R_c+P(A)}\right)^k \tag{5}$$

This expression can also be obtained formally.

P(k accidents)=

$$\sum_{N=k}^{\infty}\binom{N}{k}\left(\frac{P(A)}{1-R_c}\right)^k \left(1 - \frac{P(A)}{1-R_c}\right)^{N-k}\binom{N+k-1}{k} R_c^M\left(1-R_c\right)^N$$

$$= \frac{P(A)^k R_c^M}{k!(M-1)!} \sum_{N=0}^{\infty} \frac{(N+L)!}{N!}\left(1 - R_c - P(A)\right)^N$$

where L= k+M-1. The summation is accomplished by differentiating X^L times the geometric series L times and using Liebnitz' rule (3). to evaluate the resultant summation.

Alert Systems

The preceeding model assumes tacidly that the crossing of the red line limits by the critical parameter immediately results in failure of the mission (the only role of the alarm system being to lessen the severity of the failure consequences). However, that being the case, the designer can always change the situation by moving the red line limits closer together. In that case the alarm systems will change into ALERT SYSTEMS and will warn mission control of an impending failure of the critical component. Assuming that an in-time alert will enable the mission control to avert the impending failure(e. g. by switching in a new unit, or by continuing the mission in a different mode (from which mode an accident may occur with probability $Q_b(t)$ after t hours of operation in that mode), then the probability of accident is,

$$P(\text{accident}) = P(A) + \int_0^T \left\{1 - \prod_{i=1}^{n} (1-E_i(t)\right\} f(t)Q_b(T-t)dt$$

$$= \int_0^T f(t)\ Q_b\ (T-t)\ dt + \int_0^T \left\{\prod (1-E_i(t)\right\} f(t)R_b(T-t)dt. \tag{6}$$

where $R_b(t) = 1-Q_b(t)$.

Theorem 7: Theorems 1 through 6 continue to be correct when the costs of Alert Systems are being optimized and equation 6 is used for estimating the probability of accident instead of equation 1. The general proceedure is thus also valid.

Conclusion

The probability of accident during a mission of length T is given when there are n alarm systems, or n alert systems, (or n fail-safe devices), in parallel monitoring a critical component. These expressions have been used to minimize the total Life Cycle Costs when conditions A and B obtain. A general proceedure is given for that case. The concept of best ordering is defined and found for some important situations. It is the best orderings, if they exist, which yield the minimum Life Cycle Costs by giving the smallest n possible. After the optimum number has been found, then the number-of-accidents pdf is of interest. This pdf is given for two different operational models.

Appendix: Proofs of The Theorems

Most of the theorems follow directly from the following lemma,

Lemma 8 When p_n is defined as in equation 2, then $0 \le (p_{n+1}-p) \le 1$. If condition A is satisfied then $p_{n+1}-p_n$ decreases to zero as n approaches infinity. If condition B is satisfied then $(p_{n+1}-p_n)$ decreases monotonically;

Proof: From equation 1,

$$p_{n+1} - p_n =$$

$$r_{n+1} \int_0^T g_n(t)\ R_{n+1}(t)\ H_{n+1}\ (a(t))f(t)\ dt\ /\ Q_c(T) \quad (7)$$

where $g_n(t) = \prod_{i=1}^{n} (1-E_i(t))$.

Thus by condition A, $(P_{n+1} - p_n) < (1-s)^n / Q_c(T)$, and this approaches zero as n approaches infinity. QED.

Proof of Theorem 1. Due to physical limitations (e.g. space constraints, volume constraints etc .) $c_1(n)$ eventually becomes monatonically increasing and approaches infinity with n. Thus $c_1(x+1)-b$ has a finite largest solution. The result follows since $(p_{n+1}-p) < 1$ then implies that $\Delta(n) > 0$ for all n greater than n_1. QED

Proof of Theorem 2. Note that by lemma 8, when condition B is satisfied then $(p_{n+1}-p_n)$ is monotonically decreasing, Thus $\Delta(x)$ has at most one zero. The fact that $c_1(n)$ eventually goes to infinity and that $(p_{n+1}-p_n)$ is less then or equal to one, assures that there is exactly one solution when $\Delta(1) < 0$. Condition B also assures that $\Delta(1) > 0$ implies that $n_0=0$.

Proof of Theorem 3. From equation 7,

$$p_{n+1}-p = rH(a)\ (1-rH(a)\ )^n,$$

because a zero quiescent failure rate implies $R(t) = 1$. Using this in 3 gives,

$$c = brH(a)\ (1-rH(a)\ )^n$$

Solving for n gives y(c). The rest follows from this and the fact that $(p_{n+1} - p_n)$ goes to zero under condition A and decreases monotonically under condition B.

Proof of Theorem 5. Label the alarm systems in an arbitrary way by the positive integers. Let W be a rearrangement of the positive integers, i.e. W = w(1), w(2), ... Note that for determining best orderings, only the initial segment

$$I(W) = w(1),\ .\ .\ .\ .\ ,\ w(\ n_o(W)\ ).\ .$$

is relevent ($n_o(W)$ is the optimum number of alarm systems when the ordering W is used). Now suppose that W is an ordering which satisfies condition B. Since $c_1(n)$ is the same for every ordering W, the theorem will be proved if we can show that no segment of length $k < n_o(W)$ can be the initial segment of some ordering J. Toward this end define J to be the ordering,

$$J = s(1),\ .\ .\ .\ s(k),\ s(k+1),\ .\ .\ .$$

Using this ordering J, define,

$$h_{s(i)}(t) = 1-E_{s(i)}(t).$$

Let $p_n(J)$ be the probability of accident when the first n alarm systems labeled by the ordering J are used. We have to show that $n_o(J) \geq n_o(W)$. But if z is any integer, then,

$$Q_c(T)\ \ p_z(J) = \int_0^T \Big\{ \prod_1^z h_{s(i)}(t) \Big\}\ f(t)dt$$

$$= \int_0^T \Big\{ \prod_1^z h_{w(i)}(t) \Big\} f(t)\ K(t)\ dt$$

$$\text{where } K(t) = \frac{h_{s(i)}.\ .\ .\ .\ h_{s(z)}}{h_{w(1)}.\ .\ .\ .\ h_{w(z)}}$$

By condition B there are at most (n-1) alarms with effectiveness greater than that of the w(n) th. alarm system. Hence there are at most n-1 alarms with $h_i(t)$ less than or equal to that of the w(n) th. alarm system. Hence there is at least one integer, $i \leq n$, so that $h_{s(i)}(t) \leq h_{w(n)}(t)$ Starting with n=z continue pairing factors in the denominator with those of the numerator, (being sure not to use the same factor twice) so that the above inequality is satisfied in each case. Thus for all integers z, $K(t) \leq 1$ and $p_z(J) \leq p_z(W)$. QED.

Proof of Theorem 7: The only difference in equation 7 will be the factor $R_h(T-t)$ and hence Lemma 8 continues to hold for alert systems. This fact carries the proof of theorems 1 through 6. QED

Acknowledgements

Thanks to professors Richard E. Barlow and Donald M. Layton for drawing my attention to the safety field, and to Mr. Arnold Bakaler for a careful reading of the manuscript and useful comments.

References

1. Bellman, Richard E., and Stuart E. Dreyfus, Applied Dynamic Programming, Princeton University Press, 1962, Chapter IV.

2. Barlow, Richard E., and Frank Proschan, Mathematical Theory of Reliability, John Wiley and Sons, Inc. 1967

3. Abramowitz, M., and I. A. Stegun(eds), Handbook of Mathematical Functions, National Bureau of Standards, Applied Mathematics Series 55, 1967.

BIOGRAPHIES

ABBOTT, Wilton R.

Dr. Abbott is a Special Assistant for System Effectiveness at the Space Systems Division, Lockheed Missiles & Space Company in Sunnyvale, California. He is responsible for assisting all SSD organizations in the solution of System Effectiveness problems and the development and implementation of System Effectiveness improvement techniques. He has been with SSD in various engineering management and scientific positions since 1957.

Dr. Abbott's previous industrial experience includes positions with North American Aviation, Inc., General Electric Company, and Remler Co., Ltd. He has held academic positions at the University of California, Iowa State College and Stanford University.

Dr. Abbott received his B.S. in Electrical Engineering from the University of California and his M.S. and Ph.D. in Electrical Engineering from Iowa State College. He holds nine patents and has published numerous articles. He is a member of IEEE and Sigma Xi.

ALBERT, John S.M.

Mr. Albert graduated in 1964 from the Borough Polytechnic, London, England, and holds the Higher National Certificate (Batchelor of Science Degree) in Mechanical Engineering.

He joined Eastern Air Lines, Engineering Division, Miami, Florida, in 1967 and is currently responsible for reliability of aircraft environmental control systems involving: review of new aircraft systems design; formulation of aircraft maintenance programs; development and approval of aircraft modifications; devising new trouble-shooting techniques; Eastern's representative on the L1011 Maintenance Review Board for Aircraft Systems.

Mr. Albert's previous experience includes testing and developing prototype systems for the B737 aircraft with the Boeing Company and research into railway axle fatigue with particular reference to the effects of stress relieving grooves in reducing shear failures with the London, England, Transport Board.

Mr. Albert is an associate member of the American Society of Mechanical Engineers.

ANDRE, William L.

Mr. Andre is a Reliability and Maintainability engineer with the United States Army Weapons Command located at Rock Island, Illinois.

In 1967 he received a Bachelor's Degree in Mechanical Engineering from Louisiana State University at Baton Rouge, LA, and in 1969 received a Master's Degree in Industrial System Engineering from Texas A&M University at College Station, Texas. He also holds the U. S. Army Certificate in Maintainability Engineering, and is a Professional Engineer registered in the state of Illinois.

Projects performed have included Reliability studies on the turret systems of the M551 Sheridan Armored Reconnaisance Airborne Assault Vehicle; a proposed inertial rounds counter for enhancing Maintainability analyses of the M551; project engineer for the development of new test stands for testing large military transmissions; Maintainability analyses of the XM140 30mm automatic gun utilized on the AH-56A (Cheyenne) Army Assault Helicopter; and member of the electronics maintenance team trained to support the (SMASH) Southeast Asia Multisensor Armament System for the Cobra Helicopter.

Mr. Andre has also published papers on the subjects of 'Maintenance Ratios' in maximizing system availability and effectiveness; and 'Sequential Probability Ratio Tests' for multiple lot type sampling plans.

He is an associate member of ASME and AOA.

ASCHER, Harold E.

Mr. Ascher received a B.S. in Physics in 1956 from the College of the City of New York and an M.S. in Operations Research from New York University in 1970. He has also taken graduate courses at the Polytechnic Institute of Brooklyn and George Washington University.

Mr. Ascher was formerly employed by the U.S. Naval Applied Science Laboratory, Brooklyn, N.Y., where his activities included the development of mathematical models for complex systems such as the AN/SQS-26 and AN/BQS-13 Sonars, analysis of repairable system failure data, and the development of ground rules and statistical procedures for reliability/maintainability demonstration testing of complex systems. He is the author of several published papers in the reliability field, including one presented at the 1971 Annual Symposium on Reliability. He is presently Head of the Systems Analysis Section, Special Studies Office, Naval Research Laboratory.

Mr. Ascher is a member of the IEEE, including its Group on Reliability, and of the Scientific Research Society of America.

ASHENDORF, George

Mr. Ashendorf is a graduate of Brooklyn College with a B.A. Degree in Science, from Columbia University, with a B.S. Degree in Chemical Engineering and an M.S. Degree in Industrial Engineering.

He has seen service with the U.S. Army Ordnance Department and with the Signal Corps at Fort Monmouth, New Jersey, as a project engineer. In 1956 he joined the Burroughs Corporation as a lead reliability engineer connected with the Atlas Program.

From 1958 to 1972 he worked for RCA and had key rolls in the BMEWS and Minuteman Reliability organizations. He was also a leader in the Integrated Logistics Systems group and was responsible for systems reliability studies concerned with communications and switching systems.

Early in 1972 he rejoined the staff at Fort Monmouth, New Jersey.

BABIN, Robert S.

Mr. Babin is currently a reliability evaluation specialist in Reliability and Safety Engineering, where he specifies criteria for and performs detailed reviews and analyses of the quality, reliability, and circuit application of electronic and electrical parts in DC-10 avionics equipment. His 25 years of professional engineering experience have encompassed engineering design, development, test, and evaluation assignments in the fields of antennas, radio communication, radar, digital computers, electronic countermeasures, missile command and control, and commercial avionics.

He received the Bachelor of Electronic Engineering degree from The George Washington University in 1947 and has done graduate work in engineering and mathematics at The University of Maryland. He is the author of several technical papers and holds two patents.

BAILEY, Harold J., Jr.

Mr. Bailey is Supervisor of Monitor Test at the Conrac Division of Conrac Corporation in Covina, California. He also heads the Customer Repair Department at Conrac, providing both warranty and non-warranty service on fielded monitors. In this dual role he shares quality and reliability responsibilities with other Conrac Departments, and he directly evaluates the effectiveness of quality assurance controls in a variety of monitors.

In the past, Mr. Bailey's work has involved design review and design approval, especially of electromechanical devices. He has held project reliability responsibility in missle, ordnance, and space programs. Selected special projects in the reliability area have included studies of techniques for effective interchange of experience between commercial and military reliability fields.

Mr. Bailey received his B.S. in Electrical Engineering from the University of Minnesota and is a registered professional engineer in the state of Minnesota. He has co-authored several papers on reliability topics for national symposiums.

BALABAN, Harold S.

Mr. Balaban is a principle statistician in the Advanced Research and Development Group of ARINC Research Corporation, Annapolis, Maryland. He specializes in the development and application of mathematical, statistical and computerized models for systems analysis.

He has published a number of papers in the fields of reliability and maintainability on subjects such as models for redundancy evaluation, Bayesian approaches for reliability demonstration, Weibull parameter estimation, and reliability and maintainability prediction by function techniques.

Mr. Balaban was a co-author of the ARINC Research report, System Effectiveness: Concepts & Analytical Techniques, January 1964, in which one of the first complete model frameworks for system effectiveness quantification was developed. He served as a task group member of the Air Force Weapon System Effectiveness Industry Advisory Committee (WSEIAC).

In a study for the Air Force, Mr. Balaban was responsible for the investigation of all aspects of maintainability-demonstration procedures. Major improvements in the theory, management, and data aspects of maintainability-demonstration were developed; these are currently being adopted for general use throughout the military.

He has a bachelors degree in statistics from the City College of New York, an M.A. in Mathematical Statistics from American University and is currently a candidate for the doctoral degree in mathematical statistics at the George Washington University.

BALDONADO, Orlando C.

Dr. Baldonado is a Project Scientist at Holmes & Narver, Inc., Anaheim, California. He has developed several computer techniques for reliability and safety calculations. He did pioneering work in the field of fault tree analysis. He has used GPSS and Fortran extensively in his work. His papers have been published in Japan, Germany, Czechoslovakia, and the United States.

BAZOVSKY, Igor, Jr.

Igor Bazovsky, Jr., was born in Bratislava, Slovakia, 1942. Lived in Germany, Italy and England prior to 1957. Educated at University of Washington (B.Sc.E.E. - 1962, M.Sc.E.E. - 1969). Graduate work: Mathematics, University of California at Los Angeles.

Experience:

Scientific programming, theoretical microwave propagation studies (University of Washington). Modeling of electromagnetic compatability, reliability modeling of sequential systems, cost effective component selection, and risk analysis for reliability warranties (Airplane Division, Boeing Co.). Optimum selection of defensive weapon mix, modeling of bomb delivery systems (Military Aircraft Division, Boeing Co.). Error propagation analysis, and instrument error analysis (H.Q., Boeing Apollo Program). Design, implementation, and monitoring of a Reliability program with prediction, design review, and reliability testing per Mil-Std-781B; manufacturing, reliability, and retrofit warranty risk analysis (Electronics Division, Northrop Corporation). Theoretical studies of risk evaluation under non-stationarity (Igor Bazovsky and Associates, Inc.)

Taught two Reliability courses (Design Reliability, and System Reliability) for Boeing Continuing Education Program and one for OJT (Introduction to Reliability Assurance). Taught one for OJT course on Statistics and Error Propagation to the Apollo Technological Staff. Six previously published papers on Reliability Warranties, Risk Analysis, Optimization, and Safety.

BEAN, Eloise E.

Ms. Bean is a Senior Associate with PRC Systems Sciences Company, an operating unit of Planning Research Corporation. She received her B. A. (with honors) and M. A. in mathematics from Southern Methodist University.

Ms. Bean is currently project manager of a continuing study to collect and analyze in-flight spacecraft data and has directed all PRC efforts in this area. She is also currently engaged in reliability modeling, development of reliability, and quality assurance requirements, preparation of program documentation, and monitoring the reliability activities of various contractors and subcontractors on a classified space project. She has managed three consecutive studies for the Kennedy Space Center to develop failure rates for the KSC ground support equipment based on field experience. The most recent of these studies also involved the assessment of software reliability.

Ms. Bean has co-authored a number of papers on the reliability of space borne and ground support equipment under actual use conditions.

BELL, Raymond

Mr. Bell was born in New York City on January 24, 1931. He received a B.A. degree in Mathematical Statistics from the City College of New York in 1952 and has taken graduate courses at the City College of New York and the University of Delaware.

During his military service he was a statistical consultant with the U. S. Army Chemical Corps at Edgewood Arsenal, Maryland. In 1955, he joined the Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland as a Mathematical Statistician. Presently, he serves as Chief of the Surface to Surface Missile Systems Section of the Reliability and Maintainability Division, Army Materiel Systems Analysis Agency, Aberdeen Proving Ground, Maryland.

As Chief of the Surface to Surface Missile Systems Section, Mr. Bell has been engaged in the life cycle evaluation of the U. S. Army Stockpile of surface to surface missile systems. More recently, Mr. Bell has been involved in reliability and maintainability analysis of the tactical wheeled vehicle fleet of the U. S. Army. He has also authored numerous papers and governmental reports in both U. S. and foreign publications. During the last five years, he has three times made presentations at NATO conferences on the reliability of U. S. Army missile systems.

BIEDENBENDER, Richard E.

Mr. Biedenbender is currently the Special Assistant for Cost/Value Engineering, Office of the Secretary of Defense (Installations and Logistics). He had served as the Staff Director, Product Assurance, Office of the Secretary of Defense (Installations & Logistics) from January 1971 to March 1972. Prior to this appointment he had served as Director, Value Engineering, Office of the Secretary of Defense (I&L). He received this appointment in 1969 after serving as Director of the DoD Value Engineering Services Office since 1963.

Prior to his transfer to the Office of the Secretary of Defense, he had been continuously associated since 1951 with the Department of the Air Force in the Comptroller, Quality Control, and Procurement Functions, specializing in quality control and reliability, and industry engineering. He is the author of a number of papers in the fields of quality assurance, reliability, and incentives.

Mr. Biedenbender was born in Middletown, Ohio, on Nov 26, 1926. He received a B.S. from Dayton University in 1950, and an M.S. from Michigan State University in 1951. He has since done additional graduate work in Industrial Engineering at Ohio State and Stanford. He is a member of the American Society for Quality Control, the Operations Research Society of America, and the Society of American Value Engineers.

BLOOMQUIST, Charles E.

Mr. Bloomquist is a Senior Associate with PRC Systems Sciences Company, an operating unit of Planning Research Corporation. He received an A. B. degree in statistics from the University of California at Berkeley in 1959.

Mr. Bloomquist is currently manager of a project to study the use of the Space Shuttle to avoid or repair anomalies or unmanned spacecraft. During his 10 years with PRC he has been continuously active in various reliability aspects of the U. S. space program both civilian and military. He has been the principal investigator in two sequential analyses of spacecraft on orbit reliability and has conducted analytical reliability studies for the OGO, ATS, GEOS, and RAE satellites. He has also developed and applied a methodology for the reliability assessment of ground support equipment components at the Kennedy Space Center.

Mr. Bloomquist is a member of the American Statistical Association and has co-authored numerous papers in the field of reliability.

BOARDMAN, Howard B.

Howard B. Boardman is currently leader of the ORTS Systems Engineering group on the AEGIS Program at RCA Moorestown, and has managed the concept and design development of the ORTS system through the competitive AEGIS contract definition phase and the current engineering test and evaluation phase. Mr. Boardman has a broad background in weapons system design and instrumentation radar system engineering, which he has accumulated over the last sixteen years at the RCA Missile and Surface Radar Division. His current activity on automatic on-line monitoring and testing systems originated during a 1964 Navy-sponsored contract to evaluate availability and performance improvements for the AN/SPY-55A/B shipboard radar (TERRIER).

Mr. Boardman, born in London, England, in 1928, received his BSEE from the University of New Hampshire in 1956 and his MSEE from the University of Pennsylvania in 1963. He is a member of IEEE and Tau Beta Pi, and has authored several papers and symposium presentations.

BOEHMER, C.B.

Mr. Boehmer currently directs Reliability and Safety Analyses and Test Program Development for the Nuclear Stage Program at McDonnell Douglas Astronautics Company in Huntington Beach, California. Mr. Boehmer was previously associated with Westinghouse Astronuclear Laboratory on the NERVA project and with Los Alamos Scientific Laboratory on the pre-NERVA nuclear rocket. In these activities he was responsible for the design and operations of the control systems and the control rooms of Test Cells A and C at the Nuclear Rocket Development Station at Jackass Flats, Nevada. Mr. Boehmer prepared the test procedures and served as alternate Chief Test Operator for the first completely successful nuclear rocket reactor test series.

Mr. Boehmer served as Lead Engineer on the initial nuclear rocket safety study and initiated many of the test programs designed to investigate nuclear rocket safety. Mr. Boehmer prepared range safety reports and performed postflight analysis for Project Vanguard while with the Martin Company at Baltimore, Maryland. He also designed control systems for advanced nuclear reactors and performed system analysis studies on nuclear-powered submarines.

Mr. Boehmer holds a BS degree in electrical engineering from Washington University (St. Louis), a MS degree in physics from Drexel University, and is a graduate of the International School of Nuclear Science and Engineering at Argonne National Laboratory.

BOERCKEL, Albert, Jr.

Albert Boerckel, Jr., Quality Control Manager, Caterpillar Tractor Co., Basic Engine Plant. Has responsibility for Quality Assurance Division; Inspection Division; and Metallurgical Division, including Metallurgical Laboratory, Heat Treat Engineering and Planning, and Chemical Processing.

Studied at Bradley University, Peoria, Illinois, and Indiana University, Bloomington, Indiana.

Has had thirty-two years experience in inspection, manufacturing, and metallurgy -- twenty-two years having been in management positions.

Three years were spent in Japan as Technical Advisor, during the formative development period, in the joint venture of Caterpillar Mitsubishi.

Last two years were spent in Caterpillar Administration Offices with responsibility for corporate quality control activities, worldwide.

BREWERTON, Francis J.

Dr. Brewerton is Associate Professor of Industrial Management at Louisiana Tech University. He received a B.S. in Mechanical and Industrial Engineering, an M.B.A. and a D.B.A. in management from Louisiana State University. He has served on the Industrial Engineering faculty at L.S.U. and on the Graduate Business faculty at the University of North Dakota before joining Louisiana Tech. He has authored several research and journal articles dealing with a variety of decision science and reliability topics. He has extensive teaching, research, and consulting experience, and holds membership in AIIE, AIDS, the Academy of Management, and ORSA.

BORBERG, Henrik V.J.

Fil. mag. (eqv. M.Sc. - B.Sc.) at the Stockholm University (SU) 1962. Fil. lic. (eqv. Ph.D. in electronics) at the S.U. 1968.

Employed by Military Electronics Laboratory (FTL) in Sweden as Head of the Section for Reliability since 1970. Consultant reliability engineer 1969 - 1970. Research engineer at FTL 1964 - 1969.

CALVIN, Thomas W.

Mr. Calvin received his B.A. degree in Chemical Engineering in 1956 from the University of Toronto, Canada, and his M.S. degree in Applied Statistics in 1962 from Rutgers, The State University, New Brunswick, New Jersey.

Presently he is employed by the IBM Corporation, Components Division, as a Staff Engineer in the Reliability Studies and Statistical Support Department where he consults in statistical and reliability analysis for Product Assurance. Previously he was Manager of Statistics for the Carborundum Company and a Statistical Engineer at the American Cyanamid Company. Prior to entering the statistical field he worked as an engineer in pulp bleaching, paper converting, process equipment design, and process control instrumentation.

CATLIN, John C., Sr.

Mr. Catlin received a B.S. degree in Mechanical Engineering in 1943 from Purdue University. Since graduation he has held responsible positions in the design of heavy equipment, logistic support, quality assurance and general management.

Since joining TVA early in 1971, Mr. Catlin has been assigned to the Inspection and Testing Branch. Previously, he held positions in the Assurance Sciences with defense contractors, including being Product Assurance Manager in Advanced Design for Boeing/Vertol. He also spent five years in charge of design for a small company manufacturing power plant equipment.

CRAIG, Ronald D.

Mr. Craig joined Motorola Inc. after graduation from Arizona State University with a B.S. degree in Electrical Engineering. While at Motorola he served as design engineer on a number of transponders and automatic checkout equipment for manned and unmanned missiles. Mr. Craig joined the Boeing Company in 1966 and served as lead design engineer in the Minuteman telemetry design group.

Mr. Craig joined the Boeing SRAM System Safety group as lead engineer in 1968. In this capacity Mr. Craig has the technical responsibility for the

conduct of all safety analyses including special and nuclear safety analyses for the SRAM program. Mr. Craig has provided technical support to a number of other Boeing Safety programs and has been a guest lecturer at System Safety Courses at the University of Washington.

de CORLIEU, Jacques

Jacques de CORLIEU graduated from Ecole Polytechnique (Paris) in 1942 and from Ecole Nationale Superieure des moteurs in 1943.

He has been with THOMSON-CSF and affiliated Companies since 1960 and he has been working on weapon systems, radras and related devices. His fields of interest are Quality Control, Reliability and theory of maintenance and logistics.

He teaches system reliability at Ecole Nationale Superieure de l'Aeronautique et de l'Espace and in several other schools.

He wrote many papers about the above subjects.

D'SA, N.

Mr. N. D'Sa is a graduate student in the Department of Industrial Engineering and Operations Research at Syracuse University. After completing his Masters in Industrial Engineering and Operations Research, he plans to obtain a Masters in Systems and Information Science.

DUHAN, Stanley

Mr. Duhan received B.S. degrees in Mechanical Engineering and Industrial Engineering from the Virginia Polytechnic Institute in 1949. Since graduation, he has held responsible positions in the design, manufacturing and Product Assurance disciplines.

He is, at present, a Materials Engineer, with the Tennessee Valley Authority performing various Quality Assurance assignments.

Prior to joining TVA, Mr. Duhan was employed by the Vertol Division of The Boeing Company as the Maintainability Group Engineer for the CH-47 (Chinook) helicopter program.

Before that he was associated for 11 years with the Lycoming Division, AVCO Corporation. His last assignment with the company was as Chief, Maintainability and Safety Engineering.

Mr. Duhan has presented papers at two prior Reliability and Maintainability Symposiums.

DYLEWSKI, T.J.

T. J. Dylewski is a senior staff scientist in the Scientific Computing Division of Lockheed Missiles and Space Company, Inc., at Sunnyvale, California. He specializes in the development of operations-research approaches to the management of a large computation center, especially regarding workload forecasting, billing, capacity measurement and improvement, scheduling, and quality assurance. He has a B.S. from Illinois Institute of Technology and an M.S. from University of Cincinnati. During World War II he served in the Pacific Theatre as an Air-Force radar-countermeasures officer. Prior to joining Lockheed, he engaged in the development of various aerospace electronic and optical systems for communication, reconnaissance, and data acquisition.

EASTERLING, Robert G.

Robert G. Easterling, born September 25, 1942 in Waukegan, Illinois, is a staff member of the Statistics and Computing Division at Sandia Laboratories, Albuquerque, New Mexico. He received his B.S. in mathematics and his M.S. and Ph.D. in statistics from Oklahoma State University in 1964, 1965, and 1967, respectively. He is a member of the American Statistical Association and has served as president of the local chapter of that organization. His activities at Sandia have included consulting and research in statistical data analysis and in the application of statistical techniques to reliability assessment. Publications appear in the Journal of the American Statistical Association, the Journal of the Royal Statistical Society, B, Technometrics, the IEEE Transactions on Reliability, Proceedings of the Ninth Reliability and Maintainability Conference, and the Proceedings of the 1971 and 1972 Annual Symposia on Reliability.

EBLE, Frank A.

Mr. Eble holds degrees of Bachelor of Architectural Engineering (Catholic University of America, Magna Cum Laude, 1950), and Master of Science in Civil Engineering (University of Pennsylvania, 1967). From 1950 to 1967 he specialized in structural design and analysis of buildings, foundations, towers, and antennas, including assignments with Philco Corporation and Page Communications Engineers before joining RCA in 1960. For the past five years, he has been associated with the Technical Assurance activity of RCA's Missile and Surface Radar Division, working in the fields of reliability and maintainability engineering. Mr. Eble is a member of the AIAA and an associate member of Sigma Xi. He is a registered Professional Engineer in Pennsylvania and the District of Columbia.

EDINGER, Raymond S.

Mr. Edinger has been staff assistant to the Product Assurance Manager in technical matters at Lockheed Missiles & Space Company, Inc. since 1968. In this capacity he represents the Product Assurance Branch in planning and coordinating the Product Assurance Committee (Missile) meetings. From 1958 to 1968 he held various other positions with the same company including Product Assurance Specialist, Staff and Senior Administrator for MSD Product Assurance, Staff Administrator for PMS Missile Test Operations, and Administrator for Aerodynamics and Thermodynamics organizations in Research and Development.

He served in the United States Navy from 1942 to 1945 as a naval aviator and from 1951 to 1958, starting as a fighter pilot

in the Korean conflict and serving successively as a helicopter flight instructor, public relations officer, and finally as Aircraft Commander, Military Air Transport Service. He was awarded the Distinguished Flying Cross, the Air Medal with three clusters, a Presidential Unit Citation, and the Navy Commendation Medal.

Between 1948 and 1951 he was employed by General Motors Corporation.

He attended various LMSC and MSD training courses, Armed Forces Information School, General Motors Institute, Bloomsbury State Teachers College, Geneva College and Todds School of Aeronautics. He is a member of the American Society for Quality Control, the U.S. Naval Institute, the Naval Order of the United States and the National Management Association.

ERICKSON, John R.

Chief, Human Engineering Applications Directorate
U.S. Army Human Engineering Laboratory
Aberdeen Proving Ground, Maryland 21005

Mr. Erickson graduated with a BSME (1951) from Case Institute of Technology. He has been with the U.S. Army Human Engineering Laboratory at Aberdeen Proving Ground, Maryland, since April 1958. As Chief of the Human Engineering Applications Directorate, he is responsible for accomplishing human factors engineering of Army material for efficiency of operation and ease of maintenance, and to conduct specific applied research on those human factors affecting tactical system performance in order to derive and establish human factors design parameters to achieve the tactical effectiveness of equipment and systems.

FAGAN, Thomas L.

Mr. Fagan is Manager of Market Development at General Electric's Reentry and Environmental Systems Division at Philadelphia. In previous assignments he has managed development planning and scheduling for ERTS and other major scientific satellite programs. His experience also includes management of software development programs, management of space vehicle safety programs, and extensive analytical experience in system effectiveness, reliability and cost studies.

Mr. Fagan is a member of the evening faculty of the Philadelphia Community College (Adjunct Professor of Mathematics) and is a frequent contributor to the technical literature. He is active in community affairs, having served as a member of the United Fund campaign committee for greater Philadelphia. He is a senior member of IEEE and a senior member of ASQC. He has served as Chairman of IEEE Group 7 (reliability) and is currently a member of the management committee of the Annual Symposium on Reliability. He holds the A.B. in mathematics from Franklin and Marshall College and M.S. in statistics from Villanova University.

FINKELSTEIN, Jay L.

Mr. Finkelstein is the Project Systems Engineer in the Project Department, Navy Space Systems Activity. He has a B.A. and B.S.M.E. from Rice University and an M.S. from the California Institute of Technology.

Mr. Finkelstein is currently responsible for systems and operations analysis of space systems. This effort includes reliability and maintainability analysis. During the past 10 years he has performed various operations and operability studies on satellite systems, and has participated in design and development programs. Mr. Finkelstein has also been Program Manager on a number of missile development programs.

Mr. Finkelstein is a member of the American Institute of Astronautics and Aeronautics, the Operations Research Society and the American Society of Mechanical Engineers, and a Fellow of the American Association for the Advancement of Science. He has written a number of reports and papers on satellite systems and optimization.

FLYNN, Michael J.

He received the M.S.E.E. Degree from Syracuse University, Syracuse, New York and the Ph.D. Degree from Purdue University, Layfayette. Indiana.

He joined the IBM Corporation in 1955. He was responsible for prototype development of the 7090 and 7094 II computing systems. Later he was engaged in planning for System 360 and in study programs for development of high-speed computing

FRAGOLA, Joseph

Mr. Fragola is currently a Systems Reliability and Maintainability engineer for the Grumman Large Space Telescope (LST) program. In this position he has been responsible for the development of the econometric models used to determine the cost benefits provided by the Space Shuttle to the LST. As a member of the Systems Reliability and Maintainability group at Grumman he worked in space advanced development prior to his assignment to the LST. In this capacity he participated in the writing of many space proposals as well as conducting research into the areas of Bayesian Reliability Analysis, and the use of the Weibull distribution for reliability tracking, growth, and prediction. His publications include several in these areas. At the conclusion of his research assignment he conducted a series of training sessions at Grumman on recent developments in reliability.

As a graduate from the Polytechnic Institute of Brooklyn, Mr. Fragola received a B. S. in Physics in 1968 and an M. S. in Physics in 1971. He is presently attending the Polytechnic Institude for post graduate studies which will culminate in a PH.D. He is a member of Sigma Xi, APS, AIAA, and IEEE.

FUKUOKA, Takuji

Mr. Fukuoka is a Sinior Engineer for Reliability Assurance of the Products in Hitachi Mito Works, Japan.

He received his B.S. degree in Electrical Engineering from Muroran Technical College, Japan in 1956.

Since joining Hitachi Mito Works in 1960, he was engaged in Design of Electrical Control Equipment of Rolling stock that involved Electronic Automatic Equipment for Electric Car.

From 1965 to the present, he has been employed in Inspection Department as Relia-

bility Engineer. He has been active in Qualification and Reliability Test of Devices for Rolling stock and Elevators.

GEIGER, Robert C.

Mr. Geiger is an Assistant Program Manager with the Space Systems Division of Lockheed Missiles and Space Co. Inc. Prior to his present assignment he was Manager of Reliability and Support Services for the Standard Agena Program where he was responsible for planning and coordinating reliability and testing activities. He has an extensive background in space and aircraft systems development and testing.

Mr. Geiger is a Fellow of the Institute of Environmental Sciences and an Associate Fellow of the American Institute for Aeronautics and Astronautics. He holds an Aeronautical Engineering degree from the University of Cincinnati and Advance Study Certificates from University of California at Berkley, and at Los Angeles, and University of Santa Clara.

GLADSTONE, Samuel R.

Samuel R. Gladstone is a principal member of the engineering staff of the Missile and Surface Radar Division of RCA. At the present time he is in the Systems Engineering activity for the AEGIS Weapon Systems. His duties include responsibility for the Weapon System level RMA analyses and trade off studies.

He received his BSEE degree (1950) and MSEE (1961) from Northeastern University. After receiving his BSEE he spent the next 4 years in the U.S. Marine Corps; the last 3 years being spent in TERRIER missile checkout and firing.

His start in industry was with the Missile Systems Division of the Raytheon Co. where he was responsible for component and special pryotechnic device test and evaluation for HAWK, SPARROW, and POLARIS (guidance computer) missiles. After 9 years he went to Electron Products, for 1 year, as a Product Assurance Manager. He next spent 5 1/2 years at the Pomona Division of General Dynamics where he was responsible for the reliability analysis activities of Standard Missile I and the Standard ARM. He has been employed at RCA for over 2 years.

GOBER, R. Wayne

Dr. Gober is Professor of Statistics and Management Science at Louisiana Tech. He received a B.S. in Chemistry, an M.S. and a Ph. D. in Business Administration from the University of Alabama. He is a member of ASA, AIDS, ACM, and TIMS.

GOEL, Amrit L.

Mr. Goel is an Associate Professor of Industrial Engineering and Operations Research and Systems and Information Science at Syracuse University. Formerly, he taught as a lecturer and as an instructor at the University of Wisconsin, Madison, where he completed his M.S. and Ph. D. degrees. He has presented papers at the national and regional meetings of ORSA, American Statistical Association, Institute of Mathematical Statistics, etc. His papers have been published in JASA, Technometrics, etc.

He is a member of ORSA, ASA, ASEE, ACM, AAAS, Sigma Xi, and Fellow of the Royal Statistical Society, England.

GOLDSHINE, G.D.

Mr. Goldshine is a graduate of Rensselaer Polytechnic Institute with a BS degree in Mechanical Engineering and he did his graduate work at the University of Southern California. He is presently enrolled in the MBA program at California State College. He was employed at North American Aviation as a Thermodynamics Engineer from 1952 to 1956, and has been employed at General Dynamics, Pomona Division, since 1956. During this time he has held technical supervision positions in Preliminary Design, Controls System Design, Systems Engineering, Guidance System Design and his present position, where he is responsible for the Product Effectiveness efforts on development designs in the area of components and electronics, mechanical integration, and computer-aided design and drafting.

GREENE, Kurt

Currently head of QRC, Incorporated, a Washington, D.C. area engineering consulting and technical services firm, Mr. Greene has a Bachelor and Masters Degree in Electrical Engineering.

Mr. Greene's professional experience includes 8 years of affiliation with the Signal Corps Engineering Laboratories, where he was responsible for the development, application and evaluation of electric components; the IT&T Labs, where he was Head of the Reliability and Test Section and directed electrical components studies and evaluations to provide reliability data to equipment design and product groups; the United States Testing Company, as Manager of the Electronic Component Division, in charge of test and evaluation engineering programs and the Astro-Electronics Division of RCA, as Engineering Leader of the Reliability Engineering Group where he was responsible for the analysis of system reliability functions, and the formulation and implementation of formal Engineering Reliability Programs on Major AED projects.

In his present position he is responsible for all technical projects of QRC, Inc. which is engaged in providing consulting and engineering support services in the design and application of electrical and mechanical equipments, electrical and electro-mechanical components, the Product Assurance Sciences, test and evaluation analyses. For the past two years, QRC, Inc. has been actively engaged in the development and application of Safety Analysis Techniques.

The author of seven technical papers, Mr. Greene is a senior member of the IEEE, Chairman of the Washington Chapter of the IEEE PMP Group, and is a member of the IEEE Reliability Group Ad Com and Chairman of its Advance Technique Committee.

GRUBBS, Frank E.

Dr. Frank E. Grubbs is Chief Operations Research Analyst of the U.S. Army Aberdeen Research and Development Center, Aberdeen Proving Ground, Md. He has held this position for the last five years, prior to which he was Associate Technical Director of the Ballistic Research Laboratories (1962-1967), Chief of the Weapon Systems (Analysis) Laboratory (1954-1962), and prior to that was Chief of a laboratory evaluating the reliability of the stockpile of munitions, to fill out his thirty-one years in the Army Ordnance. During his career, Dr. Grubbs has been engaged primarily in statistical and operations research type work for the Army, especially statistical research on outlying observations, precision and accuracy, sampling inspection, reliability and life-testing, sample size determination, weapons systems evaluation, probability of hitting problems, target damage models, analysis of PERT networks and combat theory. Some of his more recent interests include confidence bounds on system reliability and new formulations of Lanchester type combat theory, which he has shown may be analyzed as a problem in Weibull reliability and life-testing theory.

Dr. Grubbs holds BS and MS Degrees in electrical engineering from Auburn University and the degrees of MA and PhD in Mathematical Statistics from the University of Michigan. He is a Founding Fellow of the American Society for Quality Control, Fellow of the American Statistical Association, Fellow of the Institute of Mathematical Statistics and Fellow of the Royal Statistical Society of London. He is an Authorized Quality Engineer and a Certified Reliability Engineer of the American Society for Quality Control.

Dr. Grubbs is author of some seventy publications on statistics, reliability, operations research, and weapons systems analysis.

In 1963, he received the Army Decoration for Exceptional Civilian Service. He is the recipient of the initial Samuel S. Wilks Memorial Medal (1964), sponsored jointly by the American Statistical Association and the Department of the Army. He won the Shewhart Medal of the American Society for Quality Control in May 1972 and was awarded both the Frank Wilcoxon prize for the best paper on practical applications and the Jack Youden prize for the best expository paper in Technometrics for 1969.

As far as Bayesian methods in reliability are concerned, Dr. Grubbs has employed all three famous techniques for reliability evaluation, including the classical approach, the fiducial method and Bayesian inference, thus covering the current important fields of interest.

GUSTAFSON, Inger A.M.

M.Sc. at the University of Stockholm 1955. Subjects: Mathematics and mathematical statistics.

In 1956 - 1970 statistician at the Material Administration of the Armed Forces. Since 1970 statistician at the Military Electronics Laboratory.

HASSLINGER, Thomas W.

Thomas W. Hasslinger was born in Gainsville, Florida on February 9, 1937. He received his B.E.E. and M.E. (Systems) from the University of Florida in 1962 and 1970, respectively.

Mr. Hasslinger has been associated with the fields of reliability and maintainability for ten years. His primary contributions have been in the areas of system analysis and design to include redundancy techniques, fail-safe techniques and design to include redundancy techniques, fail-safe techniques, operability assessment, maintenance concepts and conceptual reliability designs. Mr. Hasslinger served as Project Engineer on a NASA study to determine automated techniques for real-time status determination of redundant equipment.

Mr. Hasslinger has been employed by Radiation Incorporated since 1963 and is currently assigned to the Systems Engineering Department of Surface Operations. He is a member of the IEEE.

HAUGEN, E.B.

Professor E. B. Haugen is a faculty member in the Department of Aerospace and Mechanical Engineering, at the University of Arizona.

His activities are research, teaching (probabilistic design and experimental stress analysis), and advising higher degree candidates. He serves on the Department Graduate Studies Committee, the Design and the Materials Studies Committees.

He was for three years co-principal investigator on ONR supported research, and is now principal investigator on DOD supported research in Probabilistic Design and materials behavior. He is Director of the University of Arizona Modern Design by Reliability Institute for professional engineers, and of the National Science Foundation supported Summer Institute for College Teachers in Probabilistic Approaches to Design.

Prior to joining the faculty of the University of Arizona in 1967, he was a Research Specialist at the Space Division of North American Aviation Co, Downey, California. He has presented a number of papers before engineering and statistical societies in the United States and Europe, and has authored "Probabilistic Approaches to Design," Wiley, New York, January 1968; translated and published in a Japanese language edition in Tokyo, April 1972.

HELLER, Robert A.

Hungarian born Dr. Robert A. Heller has received his engineering education at Columbia University where he has earned a B.S. (1951) and M.S. (1953) in Civil Engineering and a Ph.D. (1958) in Engineering Mechanics. He is the author and coauthor of numerous papers on fatigue and reliability of aircraft structural materials, of a book on structures and of several educational films on mechanics of materials. Currently, he is Professor of Engineering Science and Mechanics at Virginia Polytechnic Institute and State University. He is a regional chairman of AIAA and committee chairman of ASTM.

HELLER, Agnes S.

The coauthor, Mrs. Agnes S. Heller, is the author's wife. Also Hungarian born, she has been educated at the City College of New York where she received a BBA (1954) degree in Business Administration. At Columbia University she earned an M.A. (1963) in Mathematical Statistics. Author and coauthor of many papers on reliability and statistics, she teaches statistics in the Department of Business Administration also at Virginia Tech. She is a member of the American Statistical Association and the American Institute of Decision Science.

HENDRICKS, Earl D.

Mr. Hendricks is the Increase Reliability of Operation Systems (IROS) Program Manager, as well as Head of the Reliability Engineering Group, Sacramento Air Materiel Area, McClellan Air Force Base, California. He is responsible for development and for implementation of the IROS Program and Reliability and Maintainability Engineering support to aircraft and ground communications electronics and meteorological defense systems.

Prior to his present assignment, Mr.Hendricks was a senior reliability engineer involved with numerous liquid rocket propulsion systems, including Apollo. He was the Supervisor, Process and Ballistic Control Group, responsible for small solid rocket motors at Aerojet General Corporation, Sacramento, California, from 1960-1968. He spent 1959-1960 as an operations research analyst and reliability engineer at Vought Aircraft, Dallas, Texas, after completing graduate courses required for an M.S.I.E. at Georgia Tech, 1958-1959. He received a B.S.I.E. at Texas Tech, Lubbock, Texas in 1958, and holds Professional Engineering License Number I-2121 from California.

HEREFORD, T. Graham

Professor of Humanities in the University of Virginia, Mr. Hereford brings to the Federal Executive Institute the perspective of a philosopher with educational background in science and technology as well as the liberal arts. A former Editorial Consultant for the National Bureau of Standards, he has also consulted with agencies of the State of Virginia on both educational and personnel issues. Mr. Hereford has served as chairman of the Humanities Division and Assistant to the Dean of engineering at the University of Virginia, has taught topics in philosophy and culture in three schools of the University and has lectured widely at other institutions; most recently at Oberlin College, Georgia Institute of Technology, Santa Clara University, and the General Theological Seminary of New York.

HESSE, John L.

John L. Hesse received his B.M.E. degree from the Polytechnic Institute of Brooklyn in 1956. Following three years with the Research Department, Revere Copper and Brass, Inc., Rome, New York, he entered the United States Army. He was assigned to the US Army Ordnance School as a faculty member from 1959-1960, and then served with the Southern European Task Force in Vicenza, Italy, until 1963. He then entered New Mexico State University under the Army Civil Schools Program, and received his MSME degree in 1965 and the degree, Doctor of Science in 1966. Subsequent assignments have taken him to Lawrence Radiation Laboratory, Livermore, California, as a Research Associate, and to the US Army SAFEGUARD System Evaluation Agency, White Sands Missile Range, New Mexico, as Chief, Effectiveness and Operational Reliability Division and later as Chief, Missile Site Radar Weapon Process Division. He is a registered Professional Engineer in the State of New Mexico, and is listed in the 1972 edition of American Men of Science. He currently holds the rank of Major, United States Army Ordnance Corps, and is the Commanding Officer, Kwajalein Field Office, SAFEGUARD System Evaluation Agency, Kwajalein, Marshall Islands.

HILMAN, Julian

Mr. Hilman is Deputy Chief Engineer for Product Assurance in the Engineering Division of Israel Aircraft Industries, Ltd. His responsibilities include the formulation of policies and procedures for Reliability, Maintainability and System Safety for all programs and has direct responsibility for Product Assurance of one aircraft program. He also lectures and provides Reliability consulting service to other companies.

Mr. Hilman has had 23 years of Engineering, Test and Reliability Management experience in aircraft, space and submarine programs as well as in component reliability. This includes 6 years with McDonnell Douglas Astronautics Company (head of Saturn S-IV B Reliability Program), and 3 years with Fairchild Semiconductor Company (manager of Reliability Evaluation for Minuteman components).

Mr. Hilman received his BSEE from Pennsylvania State University in February 1950 and has taken graduate courses in electronics, computer design and statistics there and at University of Pennsylvania. He is a member of IEEE, ASQC, ISQC and AIAA.

HILTON, Robert E.

Mr. Hilton came to Radiation in September 1968 as a Senior Engineer. Since that time he has been Maintainability Subtask Supervisor for the Versatile Avionic Shop Test (VAST) project and several airborne avionics programs. He has been instrumental in the development of computer programs for performance of analysis, predictions, allocations, and repair versus discard decisions. These assignments have included cost estimation, establishment of the maintainability program, performance of maintainability analysis and predictions, and responsibility for cost and schedule.

Previously, Mr. Hilton provided maintainability engineering services for Westinghouse Electric Corporation's Defense and Space Center in Baltimore, where he was involved with the AN/AWG-10 airborne radar project. He is experienced in analyzing subsystems to determine critical parameters and maintenance philosophy, including definitions of test equipment requirements.

Immediately after receiving his engineering degree, Mr. Hilton was a relay engineer for the Florida Power and Light Company, working in the areas of protective relaying, supervisory equipment, and current carrier communications.

Subsequently, he went to Chrysler Corporation's Space Division at Cape Kennedy, where he was assigned to launch site activation and checkout on the Saturn-Apollo program. He also had experience in static and rotating power sources.

Mr. Hilton is continuing his education in the computer programming field by attending the University of Michigan engineering summer conference course, "Discrete Systems Simulation Using GPSS/360."

Education: Clemson University, B.S.E.E.
University of Maryland
Florida Institute of Technology

HOLLISTER, Loren A.

Mr. Hollister is Supervisor of Inspection at the Conrac Division of Conrac Corporation in Covina, California. He is responsible for incoming, in-process, and end-item inspection of TV monitors and data display terminals. In addition, he is actively engaged in supplier selection and performance monitoring.

Mr. Hollister's previous experience includes quality engineering and supervisory responsibility in the design, fabrication, and testing of precision electomechanical assemblies. He has had similar responsibilities in both military and commercial uses of complex digital control systems.

Innovations to his credit are development and implementation of new systems for quality assurance at the various facilities where he has worked.

JACKS, Herbert G.

Mr. Jacks is a Staff Engineer in the Reliability Assurance Department at Singer-Librascope in Glendale, California. He received a Bachelor of Science degree in Engineering Physics from the University of Tennessee in 1952. His responsibilities include planning reliability and maintainability programs, conducting advanced reliability, maintainability and system effectiveness studies, and developing specific responses to customer requirements. Before joining Singer-Librascope he was with the Boeing Company where he was responsible for reliability reviews of electronic and electrical circuits. Prior to that he was with the Federal Aviation Agency where he conducted propagation and logistics studies to determine best location of radio communications stations in Alaska. Mr. Jacks is the author of a number of papers on reliability and related disciples; received the National Reliability Award in 1960; is a member of Phi Kappa Phi, IEEE and the American Ordnance Association.

KAGEY, Karen Steel

Karen Steel Kagey, M.D. graduated from New York University College of Medicine in 1960, and during Medical Residency and Cardiology Fellowship at Hartford Hospital in Connecticut became interested in many aspects of medical instrumentation use. In 1967 she joined the staff of the Peter Bent Brigham Hospital as Associate Director of Surgical Intensive Care and received appointment as Clinical Assistant In Surgery, Harvard Medical School, Boston, Massachusetts.

Daily work includes frequent contact with equipment in use, and the staff who must interact with it. The problems discussed in this paper are from first hand experience. Electrical Safety is one area of growing concern, and the Subcommittee for Electric/Electronic Appliances of which she is Chairman has become very familiar with the problems of evaluating proposed new purchases beginning with what questions to ask, and progressing to how to get answers and changes in the equipment if necessary. Dr. Kagey is active in the Boston Patient Safety Committee, the Massachusetts Hospital Association Committee on Hospital Safety and Boston Chapter of the IEEE Group on Engineering in Medicine and Biology. She is also Medical Co-chairman of the Association for Advancement of Medical Instrumentation Standards Subcommittee on Electrical Safety and member of the National Fire Protection Association Code-Making Panel No. 17 of the National Electrical Code.

KAO, John H.K.

Professor of Industrial Engineering, Department of Industrial Engineering and Operations Research, New York University, Bronx, New York 10453.

He received his B.S. in Mechanical Engineering from National Central University, M.S. in Industrial Engineering and D. Eng. Sc. both from Columbia University.

Dr. Kao is a naturalized U.S. citizen and formerly served as the engineer in charge of Purchasing and Specifications at the official agency of the Republic of China in New York City. He also has served as consultant to the U.S. Army Signal Corps and many industrial and aerospace firms on system and component reliability problems, among which are Bell Telephone Labs, Corning Glass, Electra, General Electric, Gulf-United, Pitney-Bowes, Pratt and Whitney Aircraft, United Nuclear and Westinghouse.

He is the author of more than 40 technical papers and reports (several are book chapters) on mechanical engineering and reliability problems. He has been conducting research through contract with the Office of Naval Research, on the statistical reliability techniques and theory. Three of the contract reports were chosen as Department of Defense documents: TR-3, TR-4, and TR-6 on sampling procedures and tables based on the Weibull distribution, available from the U.S. Government Printing Office.

He is the co-winner (with H. P. Goode) of the 1962-3 ASQC Electronics Award for significant contributions in the technical area of Reliability and Quality Control.

He is a member of the honory societies: Alpha Pi Mu, Phi Tau Phi and Sigma Xi, and the following professional societies: AAAS, AAUP, ASA, ORSA, senior member of ASQC, and a past member of the board of directors and vice president of the Chinese Institute of Engineers, New York, Inc.

KOHISA, T.

T. Kohisa received the B. E. degree in electronics in 1962 from Electro comunication University, Tokyo, Japan. Since 1962, he has been engaged in the field of the reliability of various semiconductor devices at Semiconductor & Integrated Circuits Division, Hitachi Ltd., Tokyo, Japan.

KOO, David Y.

Mr. Koo is a reliability engineer at Aiken Industry's Astro Communication Laboratory Division. He has been closely associated with the reliability programs for the UHF/VHF/HF receivers used in SEA WING and SEA MOUNT systems. His professional background in communication electronics also included design, technical and engineering writing, and documentation. He received his BSEE in 1956 from Case Institute of Technology and his mathematical statistician certificate in 1972 from USDA Graduate School.

KOYAMA, S.

S. Koyama received the M. S. degree in nuclear physics in 1971 from Waseda University, Tokyo, Japan. Since 1971, he has been working in the group of the reliability at Semiconductor & Integrated Circuits Division, Hitachi Ltd., Tokyo, Japan. He is a member of Physical Society of Japan.

LUTZWEIT, Walter F.

Walter F. Lutzweit, Systems Safety Specialist, Reliability and Product Safety Engineering Department, Missile Systems Division, Lockheed Missiles and Space Company Inc. (LMSC), Sunnyvale, California, has a major function in all safety at Lockheed. He was the lead engineer for the Poseidon technical manuals safety analysis, and he performed the Poseidon flight control subsystem hazard analysis-design and procedure preventive measures that were required for Poseidon. He implemented the product safety inspection program at Sunnyvale that assures safety inputs into the manufacturing system, the use of safety checklists for hazardous operations, and safety attributes for inspection verification. During his ten years at LMSC, Mr. Lutzweit has contributed to all areas of Product Assurance on Air Force Agena satellite and NASA programs in both line and staff assignments. He is a member of the LMSC Medical Emergency Rescue Corps and has received the Lockheed Certificate of Appreciation, Cost Improvement and Zero Defects Awards. Prior experience was fourteen years of research and development with eastern business machine industries. He has several U.S. Patent/Invention records and has represented companies to the U. L. Inc. Chicago Laboratory for product acceptance. Mr. Lutzweit received his B.S. degree in Physics from the University of Dayton, Ohio, in 1950 and has since completed studies at Ohio State University, as well as at the Lincoln University School of Law, and San Jose State College in California. A senior member in the IEEE since 1965, he participated in the Eleventh National Symposium on Reliability and Quality Control.

MARTIN, R.E.

Mr. Martin received a BS degree in Electrical Engineering from Columbia University in 1943 and served in the U. S. Navy as an Electronics Officer. In 1947 he joined the Naval Research Laboratory working on development of high-power transmitting and switch tubes; he then transferred to the Naval Material Laboratory, where he was in charge first of the Communication and Power Tube Section and then of the Semiconductor Devices Section. In 1963 he joined the Pomona Division of General Dynamics, where he has held the positions of Section Head of the Components, Specifications, and Standards Section and of the Drafting, Documentation, and Release Section.

MASTERSON, Robert J.

Mr. Masterson received a Bachelor of Science degree in Aeronautical Engineering from the University of Michigan. He has 21 years experience in the Aerospace Industry - 17 of which have been in Reliability.

Mr. Masterson joined TRW Systems Group in 1965 and is presently responsible for the Reliability of Mechanical and Electro-mechanical hardware in the Defense Space Systems Division. He has held similar positions on other programs and also, performed a variety of tasks on study contracts and proposals which utilized his background in the fields of Reliability, Maintainability, Availability, Operations Research, and Systems Engineering.

Prior to joining TRW Systems Group, Mr. Masterson was associated with the Denver Division of the Martin-Marietta Corporation where he was a Reliability Project Engineer in the Advanced Programs Department. Mr. Masterson's earliest experience was as Design Engineer. He first became associated with Reliability at Bell Aircraft where he was the head of the Reliability Engineering Unit.

MATTESON, Thomas D.

Mr. Matteson, who is currently Director - Maintenance Analysis for United Air Lines, has been associated with their Maintenance Operations Division for the past 12 years. Previously he was associated with Pan American World Airways. A lecturer at the University of California, Berkeley, and at the Aero Data Reliability Workshops on Reliability Analysis and Reliability Information Systems, he has written and delivered numerous papers on these subjects and on Maintenance Program Design. He is currently Chairman of the AIAA Technical Committee on Systems Effectiveness and Safety and a past Chairman of the Steering Committee for the former Annual Reliability and Maintainability Conference. He holds a MBA in Management from New York University, and a BSAE from the University of Minnesota.

MC COOL, John I.

John I. McCool was born in Philadelphia, Pa. on February 21, 1936. He received the B.S. and M.S. degrees in mechanical engineering from Drexel Institute of Technology, Philadelphia, Pa., in 1959 and 1962, respectively.

He joined the Research Laboratory of SKF Industries, Inc., King of Prussia, Pa., in 1959 and has worked in the areas of life testing and design of experiments. He is presently Supervisor of the Physics Section. Since 1965 he has been a part time Lecturer in the design of experiments and in operations research at Pennsylvania State University,

King of Prussia Graduate Center. He has had papers published in numerous technical journals.

MC MAHON, Donald J.

Donald J. McMahon is a Senior Research Statistician at the Research Center of Allegheny Ludlum Industries, Inc. He is responsible for the application of statistical and mathematical techniques to industrial situations and for the development of technical information systems.

Prior to joining Allegheny Ludlum, he was employed at the Research Center of the United States Steel Corporation. He teaches statistics at the New Kensington Campus of the Pennsylvania State University.

He received a B. S. in Engineering Science from The Pennsylvania State University in 1962 and a M. S. in Statistics from Purdue in 1964.

He currently is Chairman of the Reliability Engineering Activity of the American Society for Metals and is a past president of the Pittsburgh Chapter of the American Statistical Association. He is also a Senior Member of ASQC.

MC NICHOLS, Roger J.

Mr. McNichols is presently an Associate Professor of Industrial Engineering at Texas A&M University. He received his B.I.E., M.Sc. and Ph.D. degrees in Industrial Engineering from the Ohio State University. He is in charge of the Maintainability Engineering Program located at Red River Army Depot, Texarkana, Texas and was instrumental in the development of this program. In addition to his work with Texas A&M, Dr. McNichols is president of McNichols, Street and Associates, Inc., consulting engineers. He has been actively engaged in the development and teaching of graduate courses in reliability and maintainability and has conducted a great deal of research in these areas and has numerous publications in these and allied areas. Dr. McNichols is a registered professional engineer and a member of numerous technical societies.

MIHRAM, G. Arthur

Dr. Mihram attended, during the period June 1957 through August 1960, the University of Oklahoma, receiving there the B.S. degree in Mathematics with Special Distinction. He then undertook graduate studies in mathematics at the Washington State University, and in mathematical statistics at the Oklahoma State University, receiving the M.S. degree in 1962. He was selected as a Fulbright Scholar, studying at the University of Sydney, Australia, and completed the requirements for the Ph.D. degree in August 1965.

As an undergraduate student, he was selected for membership by PHI ETA SIGMA, PI MU EPSILON, and PHI BETA KAPPA honorary fraternities. As a graduate student, he was twice selected as a N.S.F. FELLOW and was chosen as a FULBRIGHT SCHOLAR to the University of Sydney (1964). Recently, he was elected to membership in the SIGMA XI chapter at the University of Pennsylvania.

Dr. Mihram's interest in the statistical aspects of reliability theory derive from his employment as a reliability analyst with the North American Aviation Corporation and the IBM Corporation. Current interests include general systems theory and the theory of scientific modelling, resulting in the publication of the title, SIMULATION: STATISTICAL FOUNDATIONS AND METHODOLOGY, by Academic Press in June, 1972.

Dr. Mihram is currently a member of the Faculty of the University of Pennsylvania, where he lectures and conducts research in probability theory, mathematical statistics, stochastic processes, time series analysis, simulation methodology, and the theory of scientific modelling.

MILLER, Robert N.

Mr. Miller is a Project Product Reliability Manager of the Reliability Department of Space Vehicles Product Assurance. In this capacity, he provides analytic and Operations Research related support to spacecraft studies, proposals, and hardware contracts in the fields of Reliability, Availability, Safety, System's Effectiveness. He has developed and applied several computerized techniques and procedures for performing system level tradeoffs among critical parameters of interest in the design and operation of satellite systems. Such studies and tradeoffs have been performed on many TRW programs, including Model 35, Pioneer F/G, SCS TDRS, 621B and APP.

Mr. Miller received his BA degree in Mathematics from Knox College and an MA degree in Mathematical Statistics from the University of California, Berkeley. He was a Baker Scholar at Knox College, was graduate Magna Cum Laude and admitted to Phi Beta Kappa. At Berkeley he was a Woodrow Wilson Fellow as well as a Teaching and Research Assistant in the Departments of Statistics and Electrical Engineering.

He is the co-author with G. E. Neuner of TRW, of a paper, "Resource Allocation for Maximum Reliability" which he presented at the 1966 National Reliability Symposium. Mr. Miller was the co-recipient of the 1966 National Reliability Award, the above mentioned paper having been selected as the best paper of the 1966 Symposium.

He has also authored other published papers in the Reliability field, among them being "System Optimization Using Dynamic Programming," and "Computerized Markov System Effectiveness Models," "Decision Theory in Reliability and Project Management" and "A Useful Test Design for Physics of Failure Investigations."

Mr. Miller has also served as a part-time Lecturer in Operations Research and Statistics at California State College Long Beach, teaching courses in Probability, Statistics and Decision Theory.

MIODUSKI, Robert E.

Mr. Mioduski was born in Nanticoke, Pennsylvania on July 30, 1933. He received a B.A. degree in Mathematics from Wilkes College in 1958 and has taken

48 hours of graduate work in Mathematical Statistics at the University of Delaware.

In 1958, he joined the Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland, as a Mathematician. Presently, he serves as Assistant Chief of the Surface to Surface Missile Systems Section of the Reliability and Maintainability Division, Army Materiel Systems Analysis Agency at Aberdeen.

Mr. Mioduski's primary interests are in the development and application of mathematical and statistical models in the areas of reliability, availability and maintainability. Many of his models have been successfully applied in the field of life cycle analysis of missile systems and land vehicles. Much of Mr. Mioduski's accomplishments has been published in numerous governmental reports.

Mr. Mioduski is a member of the American Statistical Association. He is also currently serving as an advisor to the Army's Tactical Vehicle Age Distribution Committee.

MORENO, Frank J.

Frank J. Moreno received the B.S. degree in mathematics from the University of Pittsburgh and the M.S. degree in Operations Research from the Florida Institute of Technology in 1961 and 1971, respectively. At the present time he is primarily concerned with the analytical treatment and modeling of communications systems from a system effectiveness viewpoint. He has also been active in the statistical analysis of field data and the design of information systems with the objective of evaluating the performance of field deployed communication systems. Previous to joining Radiation, Inc., in 1968, he worked in the areas of Operations Research and Mathematical Statistics and was mainly involved in the mathematical modeling and evaluation of weapons systems for the Department of the Army.

NAKAMURA, Yoshihiko

Yoshihiko Nakamura was born in Hokkaido, Japan on 12 July 1936. He received the B.S. degree in electrical engineering from the Hokkaido University in 1959. He joined Hitachi Works of Hitachi, Ltd. in 1959. For a few years, he was in charge of electromagnetic devices development. In 1961 he initiated proximity switches development. From 1965 to 1972, he was in charge of developing and designing many kinds of solid state devices for industrial use such as operational amplifier circuit unit, logic circuit unit, gate pulse generator and so forth. During this period, he also served in development of reliability program for solid state devices. In 1969 he transferred to Omika Works of Hitachi, Ltd. which separated from Hitachi Works and became independent. For a last few years he had worked for development and design of solid state devices applying integrated circuits for industrial use. From June 1972 he has been a senior engineer in computer control hardware engineering.

NANDA, P.

Dr. P. Nanda is an Assistant Professor of Industrial Engineering, Operations Research and Systems and Information Science at Syracuse University. His Ph.D. was from the University of Wisconsin in the area of Integer programming. His research interests include mathematical programming, reliability and maintainability, and transportation systems. He is author of a paper to appear in Management Science.

NERI, Lewis

Mr. Neri is Chief of the Reliability and Maintainability Division, Directorate for Product Assurance, U.S. Army Aviation Systems Command, St. Louis, Missouri.

He was graduated from the University of Missouri at Rolla, Missouri, 1970. In 1971, he received his Masters Degree in Engineering Management and currently is working toward his Doctoral Degree. He has been enrolled in the University of Missouri - Rolla, since April 1972.

He is now living in Pacific, Missouri. Since 1958 he has been engaged in engineering projects of various types. Responsibilities assumed have been a bridge and highway project engineer, design and development engineer on the Gemini Spacecraft project for McDonnell-Douglas, Army Project Engineer on the Army's first Attack Helicopter (AH-1G) and currently Supervisor Aerospace Engineer of the Reliability and Maintainability Division mentioned earlier. He is a licensed pilot, registered professional engineer and land surveyor.

Mr. Neri is a member of the National and Missouri Society of Professional Engineers, Missouri Association of Registered Land Surveyors, American Helicopter Society and The Army Aviation Association of America.

NOGITA, Shunsuke

Mr. Shunsuke Nogita received his B.S. in chemical engineering from Tokyo Institute of Technology. Since he joined Hitachi, Ltd. in 1959, he has been engaged in research of process control and optimization of chemical plants. His works on digital simulation of NH_3 converter, analysis and simulation of polyethylene reactor, and optimal design of olefine distillation towers were applied to chemical industry. His interests in reliability engineering resulted from the experience of process data analyses.

From fall of 1968, he spent one year's research life in Department of Chemical Engineering, Northwestern University, U.S.A.. He is a senior researcher of Hitachi Research Laboratory, and a member of the Society of Chemical Engineers, Japan.

O'LEARY, William J.

William J. O'Leary is currently leader of AEGIS system reliability, safety and standardization in the RCA AEGIS Systems Engineering group. He has managed the availability concept and design development for AEGIS through the contract definition phase and the current engineering development phase. Mr. O'Leary has also had responsibility for the application of effectiveness, system reliability, and life cycle cost techniques to such projects as Mallard, BMEWS, TERRIER, and AN/FPS-95. Before joining RCA, Mr. O'Leary was Quality Control Manager for the semiconductor facility for the Bendix Corp., Red Bank Division.

Mr. O'Leary received the BSEE and the MBA in production management from Columbia University, and the MSEE from Drexel. He is a member of the Tau Beta Pi, Alpha Kappa Psi, and ORSA.

OLIVERI, William

Mr. Oliveri is Senior Reliability Engineer in the Electromagnetic Systems Division of Raytheon Corp., where he is responsible for implementing reliability and maintainability programs for airborne transmitter jamming systems. Formerly with United Air Lines, Mr. Oliveri conducted remote and central site tests of the nationwide UNIMATIC mass communications system, comprising the real time UNIVAC 1108 system and many of its demand mode subsystems. Mr. Oliveri also wrote and successfully tested numerous FORTRAN V batch programs and their associated EXEC 8 control language statements for operating in a multiprogramming environment. These programs included a comprehensive library of mathematical and statistical subroutines and function subprograms, as well as a very general reliability/maintainability measurement report generator. Mr. Oliveri has also served as the head of the maintainability unit of Aeroject General's Electronics Division and has held reliability engineering positions on a wide variety of DOD & NASA programs such as S-3A, A-NEW, LANCE, Apollo Service Module and Saturn S-1C Booster.

Mr. Oliveri received his B.A. in Psychometrics at the University of Redlands, his M.A. in Mathematics at the University of Texas, and has pursued graduate studies in mathematics at Trinity, Georgetown, and Washington Universities. He has also taken graduate studies in physics at Incarnate Word College.

OLSEN, Alan K.

Mr. Olsen is the Chief of the Quality Management Division, San Antonio Air Materiel Area, Kelly Air Force Base, Texas. He is responsible for developing and implementing the Quality Assurance Program for the Directorate of Distribution and overseeing the AMA Inventory Control Program.

Prior to his recent reassignment, Mr. Olsen was responsible for the Reliability Engineering Branch involved in the development and application of the Reliability and Increased Reliability of Operational Systems Program. During this time, his branch was involved in the development of several major scientific data systems for the Air Force. He held this position from 1966 until June 1972. Prior to that time, Mr. Olsen was assigned as Chief of the Mechanical and Fluid Systems Section.

Mr. Olsen has held a Bachelor of Science Degree in Engineering from North Dakota State University, Fargo, North Dakota, since 1957 and received a Master of Science in Mechanical Engineering Degree in the same school in 1958. He received a Masters Degree in Public Administration from Harvard University, Cambridge, Massachusetts, in 1968 and has done further graduate work in the field of statistics at St. Mary's University in San Antonio, Texas.

ORLEANS, Beatrice S.

Beatrice S. Orleans holds a B.A. in mathematics and statistics from Hunter College and a M.S. in statistics from Columbia University. She has also completed graduate courses at Columbia, George Washington and American Universities. Prior to her present position, she was employed by the International Statistical Bureau as an economic statistician, the Educational Testing Service as a research assistant, the Air Force Human Resources Research Laboratory as an aviation psychologist and CNO Aviation Plans and Programs as a mathematical statistician. She came to the Bureau of Ships in the Statistical Engineering Branch, first in the Statistical Quality Control Section, then in the Design of Experiments Section. For the past nine years she has been Head of the Branch which also includes the Reliability Section.

She has also taught a number of sessions of a course in the Introduction to Statistical Inference for the Navy, its laboratories, and the Army.

PALKUTI, Leslie J.

Naval Research Laboratory
1966-1972, Research Electrical Engineer, Battelle, Columbus Laboratories
B.S., Electrical Engineering (1966), The Ohio State University
M.S., Electrical Engineering (1966), The Ohio State University
Ph.D., Electrical Engineering (1971), The Ohio State University

Dr. Palkuti is engaged in research concerned with semiconductor devices, their response to severe environments, studies of construction techniques, and ways to improve devices. He has recently completed studies of the lateral PNP transistor structures used in integrated circuit operational amplifiers and how these devices can be modeled. He also studied how their stability can be improved, especially to tolerate exposure to nuclear radiation.

He has prepared sections of the handbooks dealing with semiconductor devices and integrated circuits. These were aimed at improving the understanding on the part of the designer regarding the limitation of these devices under severe environments such as nuclear radiation, electromagnetic pulses, and high temperature. Understanding fabrication techniques is extremely important in these efforts.

He was involved in a research project for NASA to evaluate the effects of space radiation on microcircuits. Characterization of 650 silicon integrated circuits was performed in the laboratory. The circuits were irradiated and the data analyzed for the causes of radiation damage. Failure analyses were conducted. Recommendations were made for the application of and improvements in the integrated circuits for this environment.

From September, 1968, to September, 1969, Dr. Palkuti held a NASA Traineeship grant and served as a Teaching Associate at the Electrical Engineering Department of The Ohio State University while completing work toward his Ph.D.

Dr. Palkuti reads and speaks German and has some capability in Hungarian. He is a member of Tau Beta Pi, Sigma Xi, and Eta Kappa Nu societies, the Ohio Society of Professional Engineers, and the IEEE. He has coauthored the following papers: "Effects of Space Radiation on Silicon Microcircuits", 1968 GOMAC Digest, Volume I (October, 1968), "A Simple and Efficient Switched-Mode Approach to Hardened Power Amplification", IEEE Transactions on Nuclear Science, NS-17

(December, 1970), and "Analysis of Radiation-Induced Degradation in Lateral PNP Transistors", presented at the IEEE Annual Conference on Nuclear and Space Radiation Effects in 1971. His master's thesis was entitled "High Efficiency Audio Amplification Using Transistors in the Switch Mode".

PARASCOS, Edward T.

Mr. Parascos joined the Con Edison Company in July 1972 as Quality Assurance and Reliability Consultant.

Prior to this Mr. Parascos was employed by CBS Laboratories for nearly seven years as System Effectiveness Manager. In this capacity he managed Reliability, Maintainability, Environmental Test, Human Factors, Safety Factors, for several Image Transmission, Acquisition and processing systems including JIFDATS, DRIPS, Compass Link and E.V.R.

With the Perkin-Elmer Corporation, he managed Reliability Programs for the LEM CO_2 Sensor, Atomic Absorption Analyzer, NAPS, and the LEM Optical Tracking System Programs.

At the American Power Jet Company he managed two electromechnical reliability study programs for the Navy.

With the Kearfott Division of Singer he managed reliability efforts on MMRBM and STAFF.

Previous to this experience, Mr. Parascos spent six years as design engineer at Ford Instrument Division of Sperry Rand Corporation and Curtiss Wright Aero Division.

Mr. Parascos holds a BME and an MME from City College of New York and is presently a Ph.D. candidate at New York University.

Mr. Parascos has published several papers on Assurance Technology subjects.

PETERSON, Kenneth L.

Mr. Peterson is Lead Engineer for the DC-10 All Weather Landing Project Group in Reliability and Safety Engineering. Since 1968 he has been responsible for the development and implementation of techniques for continuously assessing system reliability requirements against inherent design capabilities, an assignment requiring a thorough knowledge of system integration and applied statistics methods. He employed applied statistics methods in systems analysis work at Pratt and Whitney Aircraft Company during 1966-68, where he was instrumental in developing several statistical reliability tests and measurement techniques.

Mr. Peterson served four years in the U.S. Navy, where he flew as navigator in carrier-based jets. He is presently a Lieutenant in the Naval Air Reserve. He received the Bachelor of Science in Mathematics degree from Wichita State University in 1966.

PIERSON, John W.

Mr. Pierson joined the System Simulation Department of Litton Data System Division as a Scientific Programmer Analyst in early 1972. He is responsible for the development of computer system simulation models for various large scale military command and control systems. Additionally, he provides performance evaluation analysis of these systems utilizing the simulation models. He has written many simulation models using high level simulation languages. He has also worked for Holmes & Narver, Inc., where he developed vehicle queing and mail flow simulation models.

PROSCHAN, Frank

Dr. Frank Proschan has been engaged in mathematical and statistical research and application for the past 30 years: from 1941-1952 for the Government, from 1952-1960 for Sylvania Electric Products, Inc., from 1960-1970 for Boeing Scientific Research Laboratories, and presently with Florida State University as Professor of Statistics. For the past fifteen years, he has been actively engaged in research in the mathematical theory of system reliability. With Dr. Richard Barlow, he has written a monograph "Mathematical Theory of Reliability" (Wiley, 1965) at the request of the Society for Industrial and Applied Mathematics.

Dr. Proschan has written about 65 papers in statistics, statistical quality control, operations research, inventory theory, and reliability; his dissertation was selected as one of the award winners of the 1959 Ford Foundation Doctoral Dissertation Competition and published by Prentice-Hall. He is a Fellow of the Institute of Mathematical Statistics, a Fellow of the American Statistical Association, and a member of the International Statistical Institute. He has been an Associate Editor of the *Annals of Mathematical Statistics* and of *Technometrics*.

Dr. Proschan received a B.S. in mathematics from the City College of New York in 1941, an M.A. in statistics from George Washington University in 1948, and a Ph.D. in statistics from Stanford University in 1959.

RAPHELSON, Morton

Mr. Raphelson received his B.E.E. Degree from Villanova University in 1950. He received his M.S. Degree in applied statistics from the same University in 1956.

Mr. Raphelson is Manager, Integrated Logistics Systems and is responsible for the Product Assurance of many programs in the Government Communications Systems Division. He is concerned with providing support on systems development in the areas of reliability, maintainability, maintenance concepts, logistics support, system safety, systems effectiveness and life-cycle cost. He is also concerned with development of analytical techniques in these areas.

He has had prior experience as Manager Product Assurance on the Minuteman Program and directed the Reliability Engineering effort on the BMEWS Project.

Prior to coming to RCA he was Supervisor Components and Reliability Section at the Burroughs Corporation, Paoli, Pennsylvania.

REHG, Virgil

Mr. Rehg is a Professor of Quantitative Methods and Statistics at the U.S. Air Force Institute of

Technology. He was associated with the Ohio State University for twelve years and spent ten years in industry prior to his present assignment.

In his present position, he is the course director for a Reliability Course presented at the School of Systems and Logistics where he also teaches statistics and operations research.

Mr. Rehg developed several simulation exercises and also has a number of articles and publications to his credit, the most recent being a textbook entitled, "Reliability, Concepts and Statistical Techniques." His formal education includes a Bachelors degree in Mathematical Statistics and a Masters degree in Business Administration, both from the St. Louis University.

RICE, Philip F.

Dr. Rice is Associate Professor of Management Science at Louisiana Tech. He received a B.S. in Electrical Engineering and a M.B.A. from the University of Arkansas, and his Ph.D. in Engineering Management from Clemson University. His teaching and research interests are in the areas of Management Science, Business Statistics, and Regional Economics. He is a member of ASA, AIDS, TIMS, Southern Economics Association and Southwestern Social Science Association.

RICHARDS, Dale O.

Dr. Richards is professor of Statistics at Brigham Young University. He received a joint Ph.D. in Statistics and Industrial Engineering from Iowa State University in 1965. This was preceded by an M.S. in Statistics at Iowa State University in 1950.

Experience includes working on contracts with Hill Air Force Base involving various reliability studies as well as serving as an operations analyst for the Iowa State University Stand-by Unit where again work with Hill Air Force Base was conducted in the area of reliability estimation and prediction studies.

Dr. Richards has been a consultant with CEIR Professional Services Division of Control Data, Collins Radio Corp., and Deseret Test Center. He has also delivered several papers at various professional meetings.

SALT, Trevor L.

Trevor L. Salt is a Design Consulting Engineer with the Group Engineering Division of the Aircraft Engine Group of the General Electric Company in Lynn, Massachusetts.

A graduate of London University England his first introduction to Aircraft Engine design was as a stress man with Rolls Royce of Derby, England in 1951. In 1959 he joined Pratt & Whitney, Canada as a senior stress engineer and moved to General Electric in Lynn in 1965. During the last five years with this company he has been employed in a design consulting role with particular recent attachment to the Preliminary Design organization.

He has presented several papers in the field of Reliability and Maintainability particularly with respect to the application of Bayes Theorem and procedures for establishment of optimum maintenance plans. In 1972 he was an invited speaker at the Reliability and Maintainability seminar sponsored by Pennsylvania State University.

SANDIN, Fredrik S.G.

Degree from Royal Institute of Technolgy 1969.

Occupation: Has been working as a reliability engineer within the Military Electronics Laboratory since that time. Prediction techniques for mechanical components as roller bearings and electromechanical relays. Current activities include systems reliability analysis.

SASSER, Gerald E., Jr.

Gerald Sasser is currently working for Intermountain Foods, a McDonald's franchise, as the General Manager of their restaurant in Provo, Utah. He received a Bachelor of Science degree in Mathematics from Brigham Young University in May, 1970. He is now working toward a Master of Science degree in Statistics at the same university. During the development of this paper, he worked as a research assistant for Dr. Dale O. Richards, the co-author.

SCHAFER, R.E.

R. E. Schafer has been associated with the Hughes Aircraft Company for the past thirteen years. At Hughes he is engaged in applied problems and research in statistical methods in reliability; in particular, mathematical models for systems, Bayesian methods of test design, and statistical methods of testing statistical hypotheses. He has published papers in I.E.E.E. Transactions on Reliability, Naval Research Logistics Quarterly, Industrial Quality Control, Biometrika, Technometrics and Operations Research. He is also a part-time faculty member at California State University, Fullerton and an Associate Editor, Technometrics.

The educational background of Mr. Schafer includes a Ph.D. degree in Statistics from Case Western Reserve University.

SCHATZ, Robert A.

Mr. Schatz first joined Westinghouse at the Transformer Division in 1951 after receiving the B.S. and B.S.E.E. degrees from Purdue University and worked in the area of design of load tap changer controls. In 1957 he was assigned to the Long-Range Major Development Group with responsibility for development of new products which will be significant commercially to the transformer industry five to ten years hence. This activity involved conducting feasibility studies; design, analyze and supervise construction and testing of models; and consultation on control and instrumentation problems. He received the M.S. degree in Control Engineering from the University of Pittsburgh in 1961.

Mr. Schatz joined the Westinghouse Astronuclear Laboratory staff as a Senior Engineer in 1963 and worked in the area of design and analysis of control equipment in support of the NRX reactor test series. In April of 1964, he was assigned to the Controls Group at the Nuclear Reactor Development Station (NDRS), Jackass Flats, Nevada, to provide analysis and operational support of the NRX-A3 test. Following this assignment, Mr. Schatz was assigned to the Systems Design and Analysis Group and appointed Fellow Engineer. He was responsible for investigations in the areas of:

Integrated circuit development,

Engine system failure - adaptability studies,

Design and analysis of the NRX-A6 reactor flow control systems,

Design and analysis of computer-supervised redundant control systems, and

Development of control and maintainability systems for SNAP-23 radioactive power source proposal.

In May of 1970, he was assigned to the Reliability Analysis Group with lead responsibility for Preliminary Design Review Data Items, probabilistic mathematical modeling, and malfunction analysis studies. He was subsequently assigned to the Systems Analysis Group with responsibility for reliability technology development and consultation, malfunction analysis studies, review of control development activity, and dynamic reactor modeling studies. In April of 1972, Mr. Schatz transferred to the Advanced Reactors Division where he is presently responsible for System Engineering and Control System Analysis Activities.

SHAW, Leonard

Leonard Shaw was born in Toledo, Ohio, on August 15, 1934. He received The B.S. degree in electrical engineering from the University of Pennsylvania, Philadelphia, in 1956, and the M.S. and Ph.D. degrees from Stanford University, Stanford, Calif., in 1957 and 1961, respectively.

Since 1960 he has been with the Department of Electrical Engineering, Polytechnic Institute of Brooklyn, Brooklyn, N.Y., where he is currently an Associate Professor. He has been a consultant to the Sperry Systems Management Division, Syosset, N.Y., and spent part of 1970 as a Visiting Professor at the Technical University of Eindhoven, The Netherlands. His research interests include stochastic control, spectral analysis, traffic control, modelling and reliability.

SHOOMAN, Martin L.

Martin L. Shooman (S'53-M'57) was born in Trenton, N.J., on February 24, 1934. He received the S.B. and S.M. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1956, and the D.E.E. degree from the Polytechnic Institute of Brooklyn, Brooklyn, N.Y., in 1961. From 1953 to 1955 he worked as a cooperative student for the General Electric Company. During 1955-1956 he held a teaching assistantship in the Department of Electrical Engineering at M.I.T. In 1956 he joined a research and development group at the Sperry Gyroscope Company, Great Neck, N.Y., where he worked on reliability mathematics and aircraft control systems. In 1958 he joined the Department of Electrical Engineering, Polytechnic Institute of Brooklyn, where he teaches graduate courses in reliability and digital computers and undergraduate courses in electrical engineering and computer science. He has done consulting for the White Sands Missile Range, RCA Astro-Electronics Division, NASA, the Sperry Gyroscope Company, and Bell Laboratories in the fields of control systems and reliability. He is author of a series of eight technical memoranda on reliability while at Sperry Gyroscope and papers on failure analysis, reliability approximations, topological reliability, spares optimization, hazard functions, and reliability modeling. He is also the author of Chapter V of Adaptive Control Systems (McGraw-Hill, 1961), the Guidance and Control Chapter of Handbook of Telemetry (McGraw-Hill, 1967), and Probabilistic Reliability: An Engineering Approach (McGraw-Hill, 1968). Dr. Shooman is presently chairman of the New York Metropolitan Chapter, and is a member of the Administrative Committee of the IEEE Reliability Group. Dr. Shooman is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi. He is co-holder of the 1967 IEEE Reliability Award for the best technical paper. He received the 1971 IEEE Reliability Award for the best technical paper.

SIMM, John H.

Mr. Simm is currently Manufacturing Manager for the Electronic Instruments Division of Beckman Instruments, Incorporated. Prior to his current assignment, Mr. Simm was Product Assurance Manager with responsibility for Reliability, Maintainability and Quality Assurance for a number of Beckman divisions including the Systems Division which served the Aerospace Industry. Prior to joining Beckman, Mr. Simm was employed by Pratt & Whitney Aircraft in East Hartford, Connecticut. He had held various positions in instrumentation and test engineering, field engineering, advanced system development, production and quality assurance engineering.

Currently he is a certified reliability engineer (ASQC) and a member for Quality Control, The Institute of Environmental Sciences and the Association for the Advancement of Medical Instrumentation. He has served on the management committee on the Annual Symposium of Reliability since 1964 and has been Electronics Division chairman of the Orange Empire section of ISQC. He has been a teacher at local colleges and delivered papers at local and regional sections of professional societies and the author of various magazines articles.

Mr. Simm attended the University of California in Los Angeles and is a native Californian now residing in the Chicago area.

SINGPURWALLA, Nozer D.

Nozer D. Singpurwalla is an Associate Professor of Operations Research at the George Washington University, Washington, D. C.

He has written over 15 papers on statistics, statistical quality control, and reliability. For the past six years he has been engaged in research and teaching in statistics, operations research, and reliability. He is co-author of a forthcoming book, Statistical Methods in Reliability and Life Testing.

SMITH, Anthony M.

Mr. Smith is currently a Technical Consultant - Operations Analysis at the General Electric Company, Reentry and Environmental Systems Division in Philadelphia, Pennsylvania. He is responsible for various hardware and

program evaluations, system analysis studies and the development of business growth opportunities. These assignments cover product lines ranging from advanced re-entry systems to industrialized housing, and technologies such as systems effectiveness, test, safety, operations research, material control and quality control.

Prior to this, he held several line management positions with responsibilities for design, reliability and test activities associated with several DOD and NASA space programs. Over his span of 17 years with the General Electric Company, he has been engaged in various aspects of aerodynamics, flight dynamics, analysis of non-nuclear defense systems, spacecraft recovery system design, testing, system engineering and reliability.

Before joining General Electric, he was an experimental aerodynamicist with the Martin Company, and nuclear power plant engineer with Westinghouse.

Mr. Smith received a BE from the Johns Hopkins University in 1953 and an MSME from Drexel University in 1961. He is the author of 19 published papers, has served as a Session Chairman and Program Committee Member in several AIAA conferences, and is a Program Vice Chairman for this Symposium. He is an Associate Fellow of AIAA and served for two years as Chairman of their Technical Committee on Systems Effectiveness and Safety.

SONTZ, Carl

Mr. Sontz is a Project Director with TRACOR, Inc. He is currently directing the development of a five year R&D program to improve integrated circuit reliability.

At TRACOR, Mr. Sontz has primarily been engaged in reliability and computer modeling studies. He was Assistant Project Director of the IVDS Program in which he directed a team of engineers responsible for monitoring the reliability and maintainability pre-production demonstration tests of the IVDS Sonar system. He served as Assistant Project Director of a program to evaluate the effect of transducer element failure on computed beam patterns of the AN/BQS-12 IDNA Sonar. He was a consultant to the ATCOMS Study Group, US Army Electronics Command, in the area of software systems an analysis.

Before joining TRACOR, Mr, Sontz was with Computer Applications Incorporated where he headed the Research and Special Projects Department. At CAI, he was a Project Manager of the program which developed a General Effectiveness Methodology (GEM), a compiler for reliability evaluation.

Mr. Sontz worked as a systems engineer at the Sperry Gyroscope Company in Carle Place, New York and at RCA's Surface Communications Laboratory. He presented the paper, "General Effectiveness Methodology," at the 30th National Meeting of the Operations Research Society of America at Durham, North Carolina, in October 1966.

Mr. Sontz received his BEE degree in 1958 from the College of the City of New York and his MEE degree in 1962 from New York University.

SORENSEN, Arthur, Jr.

Arthur Sorensen Jr. received B.S. and M.S. degrees in Mechanical Engineering from the University of Wisconsin in 1952 and 1955, respectively, and a Ph.D. in Theoretical and Applied Mechanics from the University of Illinois in 1965.

His industrial experience includes technical assignments as Engineer in the Research Department at Allis Chalmers from 1954-1957, Project Engineer in the Environmental Laboratory at AC Electronics from 1957-1959, and Senior Project Engineer in the Inertial Instruments Department of the same company from 1963-1966.

He was appointed to the faculty at the University of Wisconsin-Milwaukee in 1966, where he is now Associate Professor of Mechanics in the College of Engineering and Applied Science. His teaching experience includes a variety of courses in linear systems, applied dynamics, mechanical vibration, engineering analysis, and random fatigue. He is the author of a dozen published reports and technical papers on metal cutting, environmental simulation, gyro evaluation, fatigue analysis, and engine vibration, which reflect his professional experience in engineering mechanics.

SPAHN, Jeffrey

Mr. Spahn is employed at Grumman Aerospace Corporation as a Systems Reliability engineer on the Large Space Telescope program. Currently he is developing a model of the satellite's systems to be used in upcoming reliability and maintainability studies.

While at Grumman he has been involved in a variety of advanced development projects, including investigating the use of Bayesian techniques with the Weibull distribution to analyze system's reliability, developing computer simulations to determine the effects of varying MTBF, MTTR, and lag on turnaround time, and conducting preliminary work in software reliability. He contributed to the reliability section of the HEAO proposal, and developed computer techniques to perform zonal maintainability analyses for the Space Shuttle proposal.

Mr. Spahn received his B.S. in Physics from Stevens Institute of Technology in 1971 and is pursuing graduate studies at the Polytechnic Institute of Brooklyn. He is a member of APS and IEEE.

SPANN, Adril C.

Mr. Spann received his B.S. in E.E. from the University of Alabama in 1960, and is a registered professional engineer in Massachusetts. He has twelve years of engineering experience, of which the past ten have been in reliability and maintainability engineering. In his current assignment as Leader of the Advanced Techniques and Evaluation Unit of the Reliability Assurance Department at Sylvania, he is responsible for surveillance of, and contribution to, the engineering techniques used by this department.

Mr. Spann has gained a diverse electronics hardware experience at Sylvania, Raytheon, Laboratory for Electronics, and RCA. His experience can be roughly divided into three phases: 1. Five years of circuit analysis, limit testing, parts application analysis,

and circuit design review. 2. Three years of varied problem-solving assignments dealing with intolerable equipment failure modes, optimization of built-in-test implementation, and formulation of optimum maintenance policies. 3. Four years of working supervisory experience, embracing the spectrum of reliability/maintainability program elements, from proposal writing through demonstration testing.

SPITLER, R.H.

R. H. Spitler has managed the scientific and administrative computation centers during the past ten years at Lockheed Missiles and Space Company, Inc., at Sunnyvale, California. He received the B.E.E. degree from George Washington University in 1951 and the M.B.A. degree from Santa Clara University in 1971. After World War II he joined the Naval Research Laboratory in Washington, D. C., where he was engaged in radar attenuation studies. Prior to his present assignment he was active in the design and operation of satellite data-reduction equipment.

STEIN, Bernard

Mr. Stein is a Certified Bio-Medical Electronics Technician and has been in the field since 1960. Presently, he is Director of the Department of Scientific and Medical Instrumentation at Charles S. Wilson Memorial Hospital in Johnson City, New York. Wilson Hospital is a 470 bed teaching hospital affiliated with the State University of New York Upstate Medical Center and is the largest hospital in Central New York State.

STERNBERG, Alexander

Mr. Sternberg is a Systems Consultant in his own business and has served as Vice President, Operations of PROVO, Inc., a company specializing in the field of product and service liability prevention. He is on the staff of the National Remodelers Association as their systems consultant and has been specializing as a business systems consultant to the specialty contracting industry.

Mr. Sternberg has had more than 20 years experience working for such firms as U.S. Army Ordnance, Sonotone Corp., General Electric Company, and R.C.A. He has been providing systems consulting services to home improvement, heavy construction equipment repair, and sundry item sales companies resolving problems associated with growth and management. His work has involved such areas as management, personnel relations, training, control systems, reliability, goal setting, incentive programs, motivation, information systems, data retrieval, planning, standards, analysis of service problems, employee performance measurement, measuring departmental costs via the P and L Statement, and other related areas. He is nationally recognized in the areas of product and service liability prevention, quality control, reliability, standards, and information systems. He has held many positions including those of line and staff both in administrative and managerial capacities and is a nationally recognized lecturer and seminar leader in the above-mentioned areas.

STEVENSON, Todd E.

Mr. Stevenson is currently a Maintainability Engineer with the U.S. Army Test and Evaluation Command at the Aberdeen Proving Ground in Maryland. He holds a Bachelor's degree in Mechanical Engineering from the University of Kansas and a Master of Engineering in Industrial Engineering from Texas A&M University. While completing his postgraduate work, Mr. Stevenson was a student in the Maintainability Engineering Program at the U.S. Army Logistics Intern Training Center in Texarkana, Texas.

SUZUKI, S.

S. Suzuki received the B. E. degree in electricity in 1957 from Kanazawa University, Ishikawa-ken, Japan. He has been also engaged in the field of the reliability of various semiconductor devices since 1957, at Semiconductor & Integrated Circuits Division, Hitachi Ltd., Tokyo, Japan.

TASHJIAN, Benjamin M.

Mr. Tashjian is supervisor of Reliability Analysis of the Safety and Licensing Department of Combustion Engineering, Inc. He is responsible for directing reliability analysis of reactor protection, control and other safety related systems. He has been employed by C-E since 1968.

Previously, he was Reliability Analysis Engineer at Hamilton Standard, Division of UAC, where he performed system reliability analysis on NASA and commercial contracts.

Prior to joining UAC, Mr. Tashjian was with the Acronetic Division of General Time Corporation and Western Union Co.

Mr. Tashjian graduated from the University of Connecticut with a B.S. Degree in Electrical Engineering in 1962. He was awarded a MBA degree in 1971 from the University of Connecticut.

Mr. Tashjian is a member of the IEEE/JCNPS, SC-5 Reliability subcommittee, and EEI Equipment Availability Task Force, ad hoc Committee for developing a Nuclear Power Plant Reliability Data Collecting System. He is a member of the NSPE and is a registered professional engineer in the state of Connecticut.

THATCHER, Richard K.

Project Leader, Engineering Physics and Electronics Division, Battelle, Columbus Laboratories
B.S., Physics and Mathematics (1964), Ohio University

B.S., Electrical Engineering (1964), Ohio University

M.S., Electrical Engineering (1970), The Ohio State University

Since joining Battelle's staff in 1964, Mr. Thatcher has studied and prepared information in two main areas. They are reliability and radiation effects on electronic component parts and systems.

In the area of reliability Mr. Thatcher participated in a program to modify reliability techniques used by the Air Force in evaluating their avionics systems. Under that program Mr. Thatcher was the codeveloper of an integrated TAbular System Reliability Analysis technique, TASRA. This technique was novel in that it was designed for use by design engineers as well as specialists in reliability. It can be applied easily to a variety of systems and can predict analytical accuracy of reliability calculations. TASRA is also useful for safety and maintenance analysis. Mr. Thatcher presented a paper on this technique at the 1971 Symposium on Reliability held in Washington, D. C., January 12-14, 1971.

In addition to this work in reliability, Mr. Thatcher has just recently completed a section of a handbook for design engineers which describes how a company should incorporate radiation effects in with their reliability, maintenance, and safety procedures. This effort is, of course, primarily of interest to the military and producers of military systems. Mr. Thatcher's work in the area of radiation effects has been to edit the TREE (Transient Radiation Effects on Electronics) Handbook, the TREE (Transient Radiation Effects on Electronics) Simulation Facilities Handbook, and coedit the TREE Preferred Procedures (Selected Electronic Parts). He has also authored the Radiation Effects Information Center's report on "Permanent Effects of Nuclear Radiation on Electronic Components" and has presented several papers in the area of radiation effects on electronics.

While studying for his undergraduate degrees, Mr. Thatcher worked parttime at Battelle in the experimental determination of effects of accelerated life testing on electronic parts, fabrication of automatic data recording equipment, report evaluation in the electronic component reliability center, and study of the effects of nuclear radiation on transistors, diodes, capacitors, resistors, and other electronic components.

He is a member of Tau Beta Pi, Eta Kappa Nu, and associate member of Sigma Xi. He was the 1970 technical program chairman for the IEEE Conference on Nuclear and Space Radiation Effects and is a registered professional engineer in the state of Ohio.

TIGER, Bernard

Mr. Tiger received the B.A. Degree in Psychology, Statistics and Mathematics from Brooklyn College and the M.S. Degree from Stevens Institute of Technology in Statistics and Electrical Engineering. He has also taken graduate courses in electrical engineering at the University of Connecticut and has pursued further study of mathematics and statistics at Rutgers University.

He has been developing mathematical models for RCA systems. The models deal with reliability, maintainability, system safety, and system effectiveness. Typical systems include the Lunar Excursion Module, weapon system computers, switching systems, command and control systems, police mobile communications, recorders, SAM D, and satellite terminals. He also designs test programs and serves as a consultant on probability and statistics, operations research, and risk analysis problems. Designed test programs including reliability, effectiveness and performance evaluation of TTL and PMOS integrated circuits. He often generates design guidelines to optimize reliability, maintainability, safety, or total life cycle cost. He is currently a senior member of the Engineering Staff, RCA Government and Commercial Systems Division, Camden, New Jersey. He has presented over 60 papers and serves as an Instructor in the RCA After-Hours Graduate Engineering Study Program.

TUȘTIN, Wayne

President — Tustin Institute of Technology, Inc.

Consultant — Vibration and Shock Testing, Measurement, Analysis and Calibration

22 E. Los Olivos Street, Santa Barbara, California 93105 Telephone (805)963-1124

Mr. Tustin is a Fellow of the Institute of Environmental Sciences and a member of the American Society of Mechanical Engineers and the Instrument Society of America. He is active on Technical Committee 50, Code 120 (Vibration Testing Procedures) of the International Electrotechnical Commission. He has lectured to the Institute of Radio Engineers, the Institute of Environmental Sciences, the American Society of Mechanical Engineers, the American Society for Quality Control and the American Trucking Association.

A partial listing of Mr. Tustin's more recent published papers follows:

A test, "Vibration and Shock Test Fixture Design," coauthored by B.J. Klee and D.V. Kimball, 1971.

A text, "Vibration and Shock Test Fixture Design," coauthored by B.J. Klee and D.V. Kimball, 1971.

A monthly feature, "Vibration Topics," in **Test Engineering and Management** resumed in 1971. It previously appeared there 1963-1967.

"Vibration Test Equipment," **Sound and Vibration**, March, 1969. Also "Design Guidelines for Vibration and Shock Test Fixtures," March, 1972.

"Vibration Measurement, Analysis and Reduction," a three-article series in **Machine Design**, commencing May 29, 1969. Also "Vibration Protection Systems," October 1, 1970. The latter was condensed into **Engineer's Digest** (Great Britain), December 1970.

"Laboratory Simulation of Transportation Shock and Vibration," Proceedings of the **Packaging Progress 1971** seminar at Rochester (N.Y.) Institute of Technology.

"A Practical Primer on Vibration Testing," **Evaluation Engineering**, November/December 1969. Reprints are available from the Institute.

Papers on sinusoidal and random vibration testing, as participant in tutorial series on dynamics at **National Meeting of the I.E.S.**, Philadelphia, April 1964; San Diego, April 1966; and St. Louis, April 1968.

"Combined Environment Testing," **27th DoD Shock and Vibration Symposium** held at El Paso, February 1959. Also "A Survey of Practical Problems Encountered in Reproducing the Captive Flight Environment by Means of Shakers and Shock Test Machines," **40th Symposium**, Fort Monroe, October 1969.

UCHIYAMA, Yoshihiro

Mr. Yoshihiro Uchiyama received the BSME degree from Tohoku University, Sendai, Japan, after which he joined Hitachi, Ltd.. For several years he was responsible for the analysis of dynamics of thermal, power station and the optimization of its

control systems. The analysis of start up dynamics and combution control of once through boiler, and the optimal design of automatic load regulation were his main works in this field.

He has participated in the development of electro-hydraulic control system for the large steam turbine, in which his main occupation is security and maintenance analysis of the system. He has authored a few technical reports on boiler and turbine control systems. He is a researcher at Hitachi Research Laboratory of Hitachi, Ltd., and is a member of the Japan Society of Mechanical Engineers.

VESELY, W.E.

I. EDUCATION

BS Degree in Physics, Case Institute of Technology, 1964; MS Degree in Nuclear Engineering; PhD Degree in Nuclear Engineering, University of Illinois, 1968

II. CURRENT POSITION AND RESPONSIBILITY

Responsible for research and development in the fields of reactor physics, reliability analysis, quality control, and general statistical and probability analysis. Also serves the role of Technical Consultant within the Company.

III. EXPERIENCE

In the reactor physics field, he has written an advanced Monte Carlo program which features a general reactor geometry, exact descriptions of neutron energy transfer, pointwise resonance treatments, and general biasing options. Dynamic core allocation is used in the program for versatility of use. The program is presently the only Monte Carlo code which is routinely used within the Company for general physics computations.

In the reliability field, he has devised a time dependent methodology for fault tree evaluations. In conjunction with this theoretical approach, he has developed a computer code package for the automatic evaluation of a fault tree. The computer package is presently being used by approximately 50 different installations and corporations around the country, including NASA, the Air Force and Army, Boeing, Honeywell, General Atomic, AVCO, Hercules, and Hughes Aircraft.

In the field of statistical analysis, he has developed a set of statistical techniques for the evaluation of failure data. These evaluations include the obtainment of failure rates, determination of abnormal environments, detection of degradations including burn-in and wear-out, and the identification of deviate component and system performance. These techniques have been incorporated in a computer program. This work is being supported by the Reliability Group within the Company and represents one of their major efforts.

Dr. Vesely serves as the major technical consultant to Reliability for statistical and quantitative analyses. He serves as a consultant to a number of outside governmental agencies and serves as a guest lecturer for a number of externally held system safety and reliability courses. Dr. Vesely supervises a group of people working on various projects and he is the thesis advisor for several graduate students. He also serves the role of Company Lecturer and has taught a number of classes for Company employees. He is a member of the American Nuclear Soceity and is a member of the scientific honoraries, Sigma Xi, Tau Beta Pi, and Phi Kappa Phi.

WEBSTER, Lee R.

Mr. Webster has over 16 years of experience in the fields of reliability and maintainability engineering and operations research and is currently System Effectiveness Manager at Radiation Division of Harris Intertype Corporation. He joined Radiation in 1966 and headed the technical staff of the Director, Reliability and Quality Assurance Department for four years. In this capacity Mr. Webster provided technical support, conducted special company-supported research projects and contributed systems analysis, reliability engineering, and operations research support to other departments throughout the organization. For nearly two years he provided operations research support to the Vice President, Radiation Systems Division, Manufacturing Operations and to the Vice President of Harris Semiconductor for integrated circuit manufacturing process yield improvement studies.

Previously, Mr. Webster had been the Reliability Engineering Manager at the Electronics and Information Systems Division of the Fairchild Hiller Corporation in Bladensburg, Maryland. At Republic Aviation Division (Electronic Products) in Farmingdale, Long Island, New York, another Fairchild Hiller subsidiary, he was Supervisor, Reliability and Quality Assurance, organizing these activities to conform with the higher level quality and reliability requirements embodied in NASA specifications NPC 200-2, -3, and -4 and NPC 250-1. The aerospace projects in which Mr. Webster participated include the F-105, TFX, FIRE re-entry test vehicle, Saturn V propellant management system, meteorological satellite system studies for Nimbus, and AEROS, Advanced Orbiting Solar Observatory, and meterological sounding, data collection and photo reconnaissance systems.

Before his employment by Republic, Mr. Webster had been a member of reliability organizations at Sperry Gyroscope Corporation in Long Island, New York. He performed design review, components engineering, packaging and structural design, and environmental testing on the B-58 prime navigation system, the AN/ALQ-27 electronic countermeasures system, and the Polaris Submarine Prime Navigation System.

Mr. Webster has been a reliability consultant to a number of companies in New York and Florida and was one of the first to pass the ASQC Quality Engineering Certification examination and is also a Certified Reliability Engineer. He is a senior member of the ASME, IEEE, SOLE, and ASQC, and is the ASME representative on the Annual Reliability and Maintainability Symposium Board of Directors and Technical Program Chairman for 1973. He has delivered over 40 technical papers at international and national symposia and has published several articles in trade journals

including Mechanical Engineering. He is currently serving Florida Institute of Technology as System Engineering Curriculum Chairman and as Adjunct Professor in the graduate school. He has also conducted numerous seminars and short courses at the national level for FIT's continuing Education Program.

Mr. Webster received his B.S.M.E. from the U.S. Merchant Marine Academy; an M.S. in Applied Mathematics from Adelphi University; an M.S. in Operations Research from New York University; and has performed other graduate studies at Columbia University and the University of Connecticut.

WELKER, Everett L.

Dr. Welker received his Ph.D in mathematical statistics from the University of Illinois in 1938, remaining on the staff where he became Associate Professor of Mathematics, responsible for work in mathematical statistics through the doctorate level.

In 1947 he joined the Bureau of Medical Economic Research of the American Medical Association where he did research in vital statistics, studying mortality trends and evaluating the impact of medical progress on the age specific death rates by cause, and also consulting on statistical aspects of other AMA research programs.

In 1952, Dr. Welker joined the Weapons System Evaluation Group as a Scientific Warfare Advisor, evaluating weapons systems in limited and total warfare for the JCS and the Secretary of Defense.

From 1957 to 1963, Dr. Welker managed the Advanced Studies Department of ARINC Research Corp. He directed a study of the reliability programs of governmental and industrial agencies in the ICBM and IRBM programs for the Defense Subcommittee of the House of Representatives Appropriation Committee. He managed the development and presentation of courses in reliability engineering and he managed the reliability studies of the Saturn space booster.

In 1963 he became manager of system effectiveness analysis programs for TEMPO, General Electric Center for Advanced Studies, managing reliability and effectiveness studies in the Apollo program and in military missile programs. Dr. Welker managed a reliability study of the Coralie stage of the Europa launch vehicle in France for the French space agency, Centre National D'Etude Spatiale. He was the modeling and statistical analyst on a long range electric power generation expansion planning study for Algeria.

In midyear 1971, Dr. Welker became Staff Scientist of the Hill AFB Engineering Office of TRW, responsible for statistical and system modeling portions of USAF reliability and aging surveillance programs. He has now transferred to TRW in Redondo Beach.

Dr. Welker belongs to Sigma Xi, Institute of Mathematical Statistics, American Institute of Aeronautics and Astronautics, and other societies.

He has presented numerous research papers in theoretical statics, reliability, maintainability and system effectiveness at national symposia and university seminars.

WESTMORELAND, Maxwell E.

Mr. Maxwell E. Westmoreland was born on Dec 18, 1940 at Woodruff, S.C. He received a B.S. in Civil Engineering in 1963 from The Citadel and a M.S. in Industrial Management in 1971 from the Georgia Institute of Technology. He has been with Headquarters, U.S. Army Materiel Command, Washington, D.C., since March 1969. As an Industrial Engineer, Reliability and Systems Assessment Division, Quality Assurance Directorate, he staff supervises the AMC System Assessment Program for aircraft, electronics, land vehicles, missiles, munitions, and weapons. Before joining Headquarters, AMC, Mr. Westmoreland was with the U.S. Army Missile Command, Huntsville, Alabama, for 4 years where he was involved in system effectiveness assessment of missile systems.

He is a member of ASQC and Professional Groups on Reliability, Aircraft, and Missiles.

WHOOLEY, J.P.

J. P. Whooley is employed in the Electric Engineering Department at Public Service Electric and Gas Company as Head of the Computer Systems Group. Mr. Whooley graduated from Manhattan College, New York City in 1960 with a B.E.E. degree, and began work at Public Service in June of that year. Since that time, he has worked on computer applications in a number of areas. In 1968 he was appointed to the Equipment Availability Task Force of Edison Electric Institute and was assigned responsibility for Research Project RP-76, a computerized data collection system that provides the electric utility industry with reliability statistics from electric generating plants. In 1970 upon successful completion of RP-76 Mr. Whooley was appointed to a three-man steering committee overseeing Research Project RP-101, a data collection system designed to gather component reliability statistics for safety systems in nuclear power plants The subject matter for his paper was drawn from his work on RP-101.

WIEBE, Henry A.

Dr. Wiebe is an assistant professor in the Engineering Management Department at the University of Missouri - Rolla. He received his B.S. in Industrial Engineering from the University of Missouri - Columbia in 1960 and his M.S. in Industrial Engineering from the same institution in 1961. His Ph.D. was received at the University of Arkansas in 1970.

Dr. Wiebe joined Bell Telephone, St. Louis, Missouri, in August 1961. In 1962, he left Bell Telephone and joined the staff of the Cost Control Department of Northern Natural Gas Company, Omaha, Nebraska. He served in various industrial engineering capacities and was supervisor of an operations research group just prior to leaving for graduate school in 1966.

Since joining the faculty at the University of Missouri - Rolla, he has spent two summers with NASA at Marshall Space Flight Center as a participant on the NASA-ASEE Summer Faculty Research Program and a third summer with the same agency under a research

contract. He has also conducted several short courses on various subjects including basic reliability concepts as part of his duties with the University.

WIRSCHING, Paul H.

Paul H. Wirsching, Associate Professor of Aerospace and Mechanical Engineering, The University of Arizona, Tucson, Arizona. Received B.S.C.E. from St. Louis University; M.S. in Engineering Science, Notre Dame University; and Ph.D. in Structural Mechanics at The University of New Mexico. Has served as a member of the Technical Staff, Hughes Aircraft Company and as an Associate Professor of Mechanical Engineering, Loyola University of Los Angeles. Professional interests include shock and vibration engineering, random vibration analysis and applications, and structural reliability.

WYNOLDS, Hans

The author is currently a Business Planner engaged in product research and analysis of new business opportunities for Lockheed Missiles & Space Company, Inc. In addition, he teaches classes in engineering and systems analysis for the University of Southern California graduate school. Mr. Wynholds previously served as a financial consultant to the Treasurer's Department of Southern California Edison Company. His interests in aviation safety and economics, however, were fostered at Lockheed's California Company in Burbank. There he worked as a Safety Analyst, and later as a Senior Marketing Analyst, on the commercial L-1011 Tristar program.

At present, Mr. Wynholds is also attending Stanford University, in the Department of Engineering-Economic Systems. He earned a Bachelor of Arts degree in Mathematics from the University of California, Riverside; and two Master of Science degrees at the University of Southern California in Systems Management and Operations Research. Relative to the Assurance Sciences, the author has frequently been invited to lecture on the subject of quantitative safety analysis by the System Safety program at USC and more recently by local colleges offering similar courses. Some of the more novel aspects of these lectures were incorporated into a paper presented at the Tenth Reliability and Maintainability Conference.

INDEX OF PAPERS

Appearing in the Proceedings of the 1972 Annual Symposium on Reliability and Maintainability
and
Three Year (1970-1972) Cumulative Author and Classified Indexes

For 1970 and 1971, The Proceedings are of the Annual Symposium on Reliability; for 1972 The Proceedings are for the Annual Symposium on Reliability and Maintainability.

There are three different listings.

1. Reference List: A complete serial listing of all the papers published in the 1972 Proceedings. Each paper is numbered sequentially, beginning with 961 (following on from the 1971 Proceedings). Each listing carries title, 1972 (for the year of the Proceedings), page numbers, authors' names, and ASQC Code numbers (i.e., descriptors).

2. Author Index: A cumulative listing for 1970-1972 of authors' and coauthors' names. The appropriate sequential serial numbers are listed for each name. It is an inverse of the Reference List.

3. Three-year Cumulative Classified Index: A complete listing of the ASQC Classified Codes.* The serial numbers of the pertinent papers are listed with each Code. It is an inverse of the Reference List.

The two Indexes include all papers from the 1970-1972 Proceedings, i.e., all papers since the 15 year Cumulative Index (Nov. 1954 to Jan. 1969) which was printed in the 1970 Symposium Proceedings. The serial numbers are listed as shown: 845-903, 1970 papers listed in 1971 Proceedings of the Annual Symposium on Reliability. 904-960, 1971 papers listed in 1972 Proceedings of the Annual Symposium on Reliability and Maintainability. 961-1038, 1972 papers listed in these Proceedings.

---

* The Codes have been extended to include
# 020 Computer use
# 890 System Safety

REFERENCE LIST

Proceedings of the 1972 Annual Symposium on Reliability and Maintainability

961. System effectiveness assurance in consumer items
1972 pp. 1-6

Parascos, Edward T.

350, 813, 870

962. Final testing major appliances
1972 pp. 7-10

Nickerson, P. E.

770, 850

963. Consumer product conformance quality control cost
1972 pp.11-19

Gilmore, Harold L.

353, 611, 630

964. A complete reliability program
1972 pp. 20-23

Muench, Joseph O.

330, 600, 810

965. Reliability modeling
1972 pp. 24-29

Jones, Harold C.

821, 831

966. Failure modes, effects and criticality analyses
1972 pp. 30-37

Jordan, Wallace E.

844

967. Data collection techniques
1972 pp. 38-43

Hahn, Richard F.

841, 864

968. The reliability engineer and the interactive terminal
1972 pp. 44-50

Lyman, Karl A.

020, 831

969. Multiple regression
1972 pp. 51-58

Dagen, Herbert

020, 543

970. System effectiveness by computer
1972 pp. 59-67

Tiger, Bernard

020, 824, 882

971. The reliability of multiplexed systems
1972 pp. 68-77

Brumback, R. A.
Kowalski, R. A.

020, 824, 831

972. A Bayesian approach to reliability: Theory and simulation
1972 pp. 78-87

Tsokos, Chris P.

020, 433, 822

973. Morgantown transportation system
1972 pp. 88-93

Griswold, John W.
Ijams, Terry E.

831, 874, 894

974. Federal dual fuel project
1972 pp. 94-98

Driver, Elwood T.
Redler, Welfred M.

831, 874, 894

975. Trends in motive power reliability
1972 pp. 99-102

Hamilton, S. Graham

831

976. Reliability in rail rapid transit and commuter cars
1972 pp. 103-110

Line, Richard H.
Cord, John M.

815, 831, 884

977. A design-sensitive maintainability-prediction technique
1972 pp. 111-115

Retterer, Bernard L.

873, 874

978. Reliability simulation of memory systems with SECC
1972 pp. 116-126

Kwon, S. K.

020, 824, 831

979. A programmed guide to R/M requirements
1972 pp. 127-136

Coppola, Anthony

554, 815, 817, 830, 871

980. Planning optimum maintainability demonstrations
1972 pp. 137-147

Wilson, Myron A.

872, 875

981. Quantitative analysis of software reliability
1972 pp. 148-157

Dickson, J. C.
Hesse, J. L.
Kientz, A. C.

824, 840

982. The reliability program of the airborne computer CK37
1972 pp. 158-165

Jiewertz, Bengt

813, 840, 850

983. A perspective on computer reliability
1972 pp. 166-173

Flynn, M. J.
Szygenda, S. A.

830, 838, 872

984. Reliability meansurements for third generation computer systems
1972 pp. 174-183

Yourdon, Edward

844, 884

985. Data utilization
1972 pp. 184-189

Jacobs, Richard M.

841, 846

986. Weibull, Bayes and entropy
1972 pp. 190-199

Simkins, D. J.

420, 433, 822

987. Reliability prediction studies of degradation-prone systems
1972 pp. 200-204

deMercado, John B.

821

988. Confidence calculations for MIL-STD-781
1972 pp. 205-212

Sumerlin, W. T.

412, 554, 823

989. System safety applications to United States Air Force programs
1972 pp. 213-217

Critchlow, David M.

890, 894

990. Quantitative safety analysis
1972 pp. 218-221

Greene, Kurt
Cinibulk, Walter

844, 892

991. Formulation of an effective safety design review for the SKYLAB program
1972 pp. 222-232

Cohen, Haggai

836, 893

992. Reliability evaluation of a heart assist system
1972 pp. 233-241

Simonaitis, Darwin F.
Anderson, Ronald T.
Kaye, Michael P.

844, 894

993. Motivation for the biomedical instrument manufacturer
1972 pp. 242-244

Micco, L. A.

890

994. The reliability engineer in the health care system
1972 pp. 245-248

Taylor, Ervin F.

810

995. Current regulatory aspects of medical devices
1972 pp. 249-250

Link, David M.

891

996. A small instrument manufacturer's experience with medical instrument reliability
1972 pp. 251-254

Rose, Harold B.

810, 891

997. Designing reliability into telephone cables
1972 pp. 255-263

Apen, J. R.

710, 812, 831, 832, 860

998. Communications satellite systems reliability
1972 pp. 264-271

Grainger, George R.
Leventhal, Abraham

820, 831

999. Reliability of the global NASCOM network
1972 pp. 272-277

Stelter, L. R.

831, 840

1000. Reliability activities for Bell System transmission systems
1972 pp. 278-288

Albert, W. G.
Bangert, J. T.
Steiff, L. H.

020, 841, 853

1001. An essay on reliability cost modeling
1972 pp. 289-292

Evans, Ralph A.

350, 630, 840

1002. A surveilliance program based on a profit motive
1972 pp. 293-299

Muller, F. W.

624, 630, 851

1003. Test cost minimization for reliability assessment
1972 300-307

Naegele, G. G.
Sellinschegg, D. W.

823, 850

1004. A reliability cost-benefit analysis study
1972 pp.308-318

Nathan, Irwin

630, 840

1005. Decision theory and cost modeling
1972 pp. 319-328

Weir, W. Thomas

414, 433, 630, 850

1006. State of art of reliability practice in the European computer market
1972 pp. 329-335

Vischi, M.

800

1007. Commercial strategy and reliability in the French automotive industry
1972 pp. 336-346

Vigier, Michel G.

800

1008. System effectiveness
1972 pp. 347-353

Corlieu, Jacques de

800

1009. Complex mission worth optimization by redundancies
1972 pp. 354-360

Leclerco, Philippe R.
Marcovici, Claude M.

020, 824

1010. Reliability decisions under uncertainty
1972 pp.361-366

Johnson, R. H.

831

1011. On reliability prediction by pattern classification
1972 pp. 367-375

Pokrovsky, Felix N.

821

1012. Truncated sequential life tests: conditional probabilities and test time
1972 pp.376-385

Stracener, Jerrell T.

200, 823

1013. Statistical test plans for maintainability demonstration
1972 pp. 386-392

Balaban, Harold S.

200, 875

1014. Reliability design for an airborne ecological system
1972 pp. 393-399

Gephart, L. S.
Balachandran, V.

020, 830

1015. A method for predicting the probability of effective equipment maintenance
1972 pp. 400-406

Siegel, Arthur I.

020, 832, 872

1016. Stochastic modeling of human performance effectiveness functions
1972 pp 407-416

Regulinski, T. L.
Askren, W. B.

425, 832

1017. A test of a basic assumption of human performance modeling
1972 pp. 417-419

Lamb, Jerry C.

832, 850

1018. Prediction of human performance
1972 pp. 420-429

Teichner, Warren H.

832

1019. Growth curves – a practical management tool
1972 pp. 430-436

Kreuze, F. J.

850, 860

1020. Design review – a tool for all seasons
1972 pp. 437-445

Liberman, David S.

836

1021. Application of decision-making models
1972 pp. 446-457

Levitt, Eugene R.

020, 414, 831

1022. NASA program decisions using reliability analysis
1972 pp. 458-464

Steinberg, Alvin

840

1023. A practical analysis of the sedentary failure rate
1972 pp. 465-473

Cherkasky, Stanley M.

840

1024. MOS/LSI reliability prediction
1972 pp. 474-478

Nichols, Elliott D.

020, 833, 840

1025. Failure mode control in plastic IC's
1972 pp. 479-491

Edfors, H. C.

833, 844

1026. Thermal cycling ratings of power transistors
1972 pp. 492-498

Lukach, V. J.
Gallace, L.
Williams, W.

833

1027. Reliable interconnections for Army avionics
1972 pp. 499-509

Spergel, Jack
Godwin, Elmer F.

833, 835

1028. MTBF prediction techniques – a comparative analysis
1972 pp. 510-519

Muglia, V. O.
Bridgers, J. E.
Cici, A. S.
Vayos, N. T.

831, 840

1029. Assessing missile system reliability
1972 pp. 520-525

Parrish, Peter P.

824, 840

1030. Reliability analysis of transmissions
1972 pp. 526-532

Mouradian, G.

824, 840

1031. Is there a reliable screen for VHF power transistors?
1972 pp. 533-540

Ingram-Cotton, J.
Sulway, D. V.
Leduc, J. P. M.

630, 815, 833

1032. Conditions related to the assumption of independence
1972 pp. 541-547

Bell, S.
Zimmer, W. J.

410, 821

1033. On combining component and system data
1972 pp. 548-553

Easterling, Robert G.

410, 820

1034. Economics of reliability for spacecraft computers
1972 pp. 554-564

Hecht, Herbert

630, 831

1035. A moment method for the calculation of a confidence interval for the failure probability of a system
1972 pp. 565-577

Murchland, John D.
Weber, Gunter G.

412, 820

1036. Development of Douglas commercial aircraft reliability programs
1972 pp. 578-585

Gilles, D. L.

800

1037. Software reliability engineering
1972 pp. 586-593

Mulick, R. B.

830, 840

1038. Reliability and ESP
1972 pp. 594-598

Maybell, J. L.

840

CUMULATIVE AUTHOR INDEX

1970–1972 Proceedings

CUMULATIVE CLASSIFIED INDEX

1970–1972 Proceedings

The ASQC Code number is followed by the Code name. The appropriate serials are listed below them. (Some Codes were not assigned to any papers.)

IEEE GROUP ON RELIABILITY

Administrative Committee

President: V.R. Monshaw, RCA, Princeton, New Jersey
Vice President/Education Development and Standards: J.J. Naresky, RADC (RC), Griffis Air Force Base, New York
Vice President/Technical Operations: C.R. Knight, ARINC Research Corporation, Annapolis, Maryland
Vice President/Publications: P.K. McElroy, Consultant, Belmont, Massachusetts
Vice President/Meetings, Chapters and Membership: H.W. Williams, Mitre Corporation, McLean, Virginia
Junior Past Chairman: H.E. Reese, Burroughs Corporation, Paoli, Pennsylvania
Treasurer: R.M. Jacobs, Newark College of Engineering, Newark, New Jersey
Secretary: D. Troxel, RCA, Camden, New Jersey

Members

J.E. Condon, NASA Headquarters, Washington, D.C.
G. Ebel, Singer-General Precision, Little Falls, New Jersey
R.A. Evans, Consultant, Durham, North Carolina
K. Greene, QRC Inc., Silver Spring, Maryland
P. Gottfried, Consultant, Silver Spring, Maryland
H.D. Hulme, Westinghouse Electric Co., Elmira, New York
R.M. Jacobs, Consultant, Livingston, New Jersey
C.R. Knight, ARINC Research Corp., Annapolis, Maryland
R.E. Kuehn, IBM Corporation, Owego., New York
P.K. McElroy, Consultant, Belmont, Massachusetts
V.R. Monshaw, RCA, Princeton, New Jersey
J.J. Naresky, RADC (RC), Griffis Air Force Base, New York
A.R. Park, Precision Instruments Co., Palo Alto, California
D.S. Peck, Bell Telephone Laboratories, Allentown, Pennsylvania
M.L. Shooman, Brooklyn Polytechnic Institute, Brooklyn, New York
M.P. Smith, Honeywell, Inc., St. Petersburg, Florida
J.W. Thomas, Vitro Laboratories, Silver Spring, Maryland
H.W. Williams, Mitre Corporation, McLean, Virginia

Chapter Chairmen

Baltimore:
Richard Kowlaski, Baltimore, Maryland
Binghamton:
J.J. Rexford, Owego, New York
Boston:
Donald L. Dawes, Lincoln, Massachusetts
Canaveral:
Jack R. Lykins, Satellite Beach, Florida
Chicago:
Charles W.N. Thompson, Glencoe, Illinois
Connecticut:
J.R. Breton, New London, Connecticut
Florida West Coast:
Robert P. Dalton, Clearwater, Florida
Los Angeles:
Robert S. Babin, Palos Verdes, California
Mohawk Valley:
George Kasouf, Utica, New York
New Jersey Coast:
Weldon V. Lane, West Long Branch, N.J.
New York, Met.:
Martin L. Shooman, Brooklyn, New York
North Jersey:
Stanley M. Cherkasky, Little Falls, N.J.
Philadelphia:
Edmund J. Westcott, Riverton, N.J.
San Francisco:
Gilmore Bowers, Menlo Park, California
Twin Cities:
Stephen Pech, Circle Pines, Minnesota
Washington, D.C.:
Wm. E. Wallace, Jr., Silver Spring, Md.

1972 AIAA TECHNICAL COMMITTEE FOR SYSTEM EFFECTIVENESS AND SAFETY

T.D. Matteson, *Chairman*, United Airlines, Inc., San Francisco, California
William M. Bland, Jr., NASA Manned Spacecraft Center, Houston, Texas
Philip Bolger, Department of Transportation, Washington, D.C.
Robert L. Bresniker, United Technology Center, Sunnyvale, California
Robert B. Bussler, Naval Air Systems Command, Washington, D.C.
E.J. Dofter, Chrysler Corp., Space Division, New Orleans, Louisiana
J. Don Marioni, System Development Corp., Santa Clara, California
James L. Maybell, McDonnell Douglas Astronautics Co., Huntington Beach, California
William S. McQuay, Jr., Bendix Corporation, Baltimore, Maryland
Allen G. Norem, Aerospace Corporation, Los Angeles, California
I. Irving Pinkel, Retired, Fair View Park, Ohio
William C. Pitts, Jr., Vought Aeronautics, Dallas, Texas
William Quinlivan, Lockheed California Company, Burbank, California
John F. Raye, Hamilton Standard, Windsor Lock, Connecticut
Harlan A. Reesing, Douglas Aircraft, Long Beach, California
William C. Schultz, Cornell Aeronautical Lab., Inc., Buffalo, New York
Paul L. Wickham, North American Rockwell Corp., Downey, California
George W. Windolph, The Boeing Company, Vertol Division, Philadelphia, Pennsylvania

AMERICAN INSTITUTE OF INDUSTRIAL ENGINEERS

President: Robert W. Lockwood, United California Bank
President-elect: Roy L. Allen, The City of Dallas. Texas
Executive Vice President: Roy R. Hunt, Crest Container Corporation
Executive Director: Jack F. Jericho, AIIE
Vice President, Finance: Clayton L. Davis, Jr., Manufacturing Data Systems Inc.
Vice President, Professional Relations and EJC Affairs: Harold T. Amrine, Purdue University
Vice President, Publications: Guy J. Bacci, II, International Harvester Co.
Vice President, Technical and Division Affairs, William A. Smith, Jr., Lehigh University
Vice President, Education and Student Affairs, Benjamin W. Niebel, Pennsylvania State University
Vice President, Region I: Eugene J. Broderick, IBM Corporation
Vice President, Region II, Edward R. Zebrowski, Rohm & Haas Company
Vice President, Region III: John R. Canada, North Carolina State University
Vice President, Region IV: Nelson C. Wall, Georgia Institute of Technology
Vice President, Region V: Nicholas Kralles, Eastman Kodak Company
Vice President, Region VI: Clarence Darrow, Jr., Federal-Mogul Corporation
Vice President, Region VII: Charles E. Clark, Jr., Ernst and Ernst
Vice President, Region VIII: V. Sam Carlile, Delco Remy Division, GMC
Vice President, Region IX: Donald J. Peterson, Western Electric Company
Vice President, Region X: Roy Kahn, Performance Investors Inc.
Vice President, Region XI: John F. Sweers, Western Publishing Co., Inc.
Vice President, Region XII: Hewitt H. Young, Arizona State University
Immediate Past President: Robert W. Newsom, Jr., Lorillard
Second Past President: Frank E. Cotton, Jr., Mississippi State University

AMERICAN SOCIETY OF MECHANICAL ENGINEERS

Aerospace Division

Chairman: Kenneth M. Foreman, Grumman Aircraft Engineering Corp., Bethpage, L.I.N.Y.
Vice Chairman: O.H. Fedor, Kennedy Space Center, NASA, Fla.

Vice Chairman: Robert F. Stoessel, Management Enterprises, Corona de Mar, Calif.

Vice Chairman: Dr. A. H. Hausrath, TRW Systems Group, San Bernardino, Calif.
Secretary: L.R. Webster, Radiation Division Harris Intertype, Melbourne, Fla.

General Committee Chairmen

Reliability: C.F. Bower, McDonnel Douglas Corporation, Huntington Beach, Calif.
Maintainability: D.D. Gregor, Northrop Corp., Aircraft Division, Hawthorne, Calif.
Systems Engineering: H.M. Kavanagh, Jet Propulsion Laboratory, Pasadena, Calif.

Ground Support & Test: R.E. Hilton, Jr., Radiation Division Harris Intertype, Melbourne, Fla.
Design & Equipment: J.C. Sanders, NASA Lewis Research Center, Cleveland, Ohio
Energy Conversion: C.H. Shinbrot, McDonnel Douglas Astronautics Co., Huntington Beach, Calif.

Crew Systems: D.C. Popma, NASA Headquarters, Washington, D.C.
Transportation: R.F.Stoessel, Management Enterprises, Inc., Corona del Mar, Calif.
Structures & Materials: Dr. H. Hurwicz, McDonnel Douglas Astronautics, Huntington Beach, Calif.

AMERICAN SOCIETY FOR QUALITY CONTROL

Electronics Division

OFFICERS

Chairman: J. Douglas Ekings, Xerox Corp., Rochester, N.Y.
General Vice Chairman: Julian Edelman, Sylvania, Needham, Mass.

Vice Chairman, Committees and Activities: Richard Deutsch, Espey Mfg., & Elect. Corp., Saratoga Springs, N.Y.
Secretary: Edward O. Shrieve, Corning Glass Works, Bluffton, Indiana
Treasurer: Dr. Ralph Evans, Evans Associates, Durham, N.C.

Junior Past Chairman: Allen Chop, Lockheed Missiles & Space Co., Sunnyvale, California
Senior Past Chairman: Naomi J. McAfee, Westinghouse Electric Corp., Baltimore, Md.

REGIONAL COUNCILORS

Region 1
Everett Milott, Sylvania Comm. Systems Div., Needham, Mass. 02194
Region 3
Richard Jacobs, Newark College of Eng'g, Newark, New Jersey 07102
Region 4
J. Fitzsimons, Montreal 248, P.Q., Canada
Region 5
John Stanley, Leeds & Northrup, North Wales, Pa. 19454

Region 6
James Whittaker, Lockheed Missiles & Space Co., Sunnyvale, California 94088
Region 7
R. Nicholsen, General Dynamics, San Diego, California 92112
Region 9
Virgil Rehg, Dayton, Ohio 45431
Region 10
V.N. Pruett, Plymouth Division, Burroughs Corp., Plymouth, Michigan 48170

Region 11
Paul D. Brown, Winston-Salem, N.C. 27107
Region 12
Donald R. Mittelstadt, UNIVAC-Div. of Sperry Rand, St. Paul, Minn. 55113
Region 14
B.W. Marguglio, Dallas, Texas 75234
Region 15
Harold Z. Snyder, Electronics Communications, Inc., St. Petersburg, Florida 33733

Reliability Division

OFFICERS

Chairman: W.E. Sawyer, Western Electric Company, Allentown, Penna.

Vice Chairman, General: H.D. Hulme, Westinghouse Electric Corp., Elmira, New York
Treasurer: A.I. Moskovitz, NASA, Washington, D.C.

Chairman-Elect: R.W. Smiley, Lockheed Missiles & Space Co., Sunnyvale, California
Secretary: E.O. Peters, Lockheed Missiles & Space Co., Sunnyvale, California

COMMITTEES

Program Chairman: J. E. Condon, Abbott Laboratories, Chicago, Illinois
Publicity & Promotions Chairman: G.W. Corley, Westinghouse Electric Corp., Pittsburgh, Penna.
Membership Committee Chairman: F. Jacobsen, Western Electric Co., Allentown, Penna.
International Correspondent Chairman: E.J. Lancaster, Lockheed Missiles & Space Co., Sunnyvale, California

Maintainability Committee Chairman: E.P. O'Connell, Performance Engineering Co., Newport, Rhode Island
IEC-TC56 Liaison Committee Chairman: J.D. Ekings, Xerox Corp., Pittsford, N.Y.
Publications Chairman: M.R. Cohen, ITT Headquarters, New York City, New York
Education Chairman: M.E. Liebman, Aerojet Solid Propulsion Co., Sacramento, Calif.

Newsletter Editor Chairman: A.I. Moskovitz, NASA, Washington, D.C.
Reliability Reporting Committee Chairman: A.S. Golant, NCR Data Processing Div., Hawthorne, California
Examining Committee Chairman: E.J. Lancaster, Lockheed Missiles & Space Co., Sunnyvale, California
Auditing Committee Chairman: T.J. Williams, Mack Trucks, Inc., Allentown, Penna.
By-Laws Committee Chairman: M. Mitchell, Western Electric Co., Allentown, Penna.

INSTITUTE OF ENVIRONMENTAL SCIENCES

President: Gerard A. Gallagher, Ecology and Environment, Inc., Buffalo, NY 14225
Executive Vice President: Frank W. Hallstein, Litton Ship Systems, Manhattan Beach, CA 90266
Fiscal Vice President: Frederick W. Rusch, North American Rockwell, Palos Verdes Estates, CA 90274

Technical Vice President: Richard T. Nichols, Lockheed Missiles & Space, San Jose, CA 95125
Vice President, Region I: Arnold Waterman, Pratt & Whitney Aircraft, East Hartford, CT 06118
Vice President, Region II: Albert E. Armstrong, Westinghouse Electric Corp., Ellicott City, MD 21043
Vice President, Region III: Frank Snidarich, Donaldson Co., Inc., Bloomington, MN 55431

Vice President, Region IV: Paul Turkheimer, Wyle Laboratories, Torrance, CA 90505
Immediate Past President: Arthur M. Dallon, Avco Corporation, Lawrence, MA 01841
Executive Director: Betty Peterson (Mrs. A.), Institute of Environmental Sciences, Mount Prospect, IL 60056
Technical Editor: Frederic K. Van Biene, Jet Propulsion Laboratory, Pasadena, CA 91103